

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 02:20:31 ; Search time 108 Seconds
(without alignments)
2019.139 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005

Sequence: 1 MTSVSAEHLGKDRPPVLVS.....HLHYRLPAADVRRPIISVRK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2005	100.0	379	2	Q84BC5 rhodococcus
2	609.5	30.4	459	2	P94171 arcanobacte
3	569	28.4	433	2	Q8GDF1 propionibac
4	497	24.8	403	2	Q89764 propionibac
5	425.5	21.2	482	2	Q84GF4 corynebacte
6	424.5	21.2	476	2	Q9RIQ0 streptomyc
7	409	20.4	525	2	Q74211 mycobacteri
8	409	20.4	525	2	AA03171 mycobacte
9	374.5	18.7	405	2	Q8SCH7 propionibac
10	373	18.6	380	2	Q9X3U7 bifidobacte
11	360.5	18.0	622	2	Q8GGD7 streptomyc
12	345	17.2	471	2	Q8RIU6 streptomyc
13	342.5	17.1	461	2	Q931L4 corynebacte
14	326	16.3	401	2	Q9R881 corynebacte
15	316	15.8	456	1	REP_STRLI
16	315.5	15.7	312	2	Q8GHF1 streptomyc
17	293.5	14.6	396	2	Q8GHF1 pseudomonas
18	277.5	13.8	396	2	Q9AGT3 rhizobium m
19	266.5	13.3	528	2	Q52220 plasmid pab
20	265	13.2	478	2	Q57477 streptomyc
21	256.5	12.8	505	2	Q8RNS3 corynebacte
22	235.5	11.7	481	2	Q83016 streptomyc
23	230.5	11.5	437	2	Q54245 streptomyc
24	227	11.3	437	1	Q52205 plasmid psb
25	214.5	10.7	389	2	P14501 streptomyc
26	194	9.7	396	2	Q84G55 corynebacte
27	169	8.4	240	2	Q69069 corynebacte
28	163.5	8.2	313	2	Q45282 streptomyc
29	161.5	8.1	261	2	Q84FM7 streptomyc
30	127	6.3	314	1	Q8GI56 thermus aqu
31	123.5	6.2	315	2	REP_LACHI
					Q56129 streptococc

32	123	6.1	334	2	Q6DRR7 streptococc
33	122.5	6.1	315	2	Q842E5 streptococc
34	122.5	6.1	315	2	Q9RNV8 streptococc
35	122.5	6.1	315	2	Q9RNV4 streptococc
36	121.5	6.1	315	2	Q9RNV5 streptococc
37	119	5.9	319	2	Q9ZNC0 streptococc
38	115	5.7	336	2	Q9ZNC0 nitrosomona
39	114.5	5.7	311	2	Q6460 synechocyst
40	114.5	5.7	311	2	Q70LGH streptococ
41	114.5	5.7	315	2	CAD97600 streptoco
42	114	5.7	314	2	Q9X9N4 streptococ
43	111.5	5.6	276	2	Q30850 streptococ
44	111.5	5.6	320	2	Q9K4S9 pseudomonas
45	111.5	5.6	320	2	O52769 streptococ
					O52776 streptococ

ALIGNMENTS

RESULT 1

Q84BC5

ID Q84BC5 PRELIMINARY; PRT; 379 AA.

AC Q84BC5; DT 01-JUN-2003 (Tremblrel. 24, Created)

DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

DE Replication protein.

GN Name=rep;

OS Rhodococcus erythropolis.

OG plasmid pAN12.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

CC Corynebacterineae; Nocardiaceae; Rhodococcus.

OX NCBI_TaxID=1833;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22718480; PubMed=12835922;

RA Kostichka K., Tao L., Bramucci M., Tomb J.-F., Nagarajan V., Cheng Q.;

RT "A small cryptic plasmid from Rhodococcus erythropolis:

characterization and utility for gene expression.";

RL Appl. Microbiol. Biotechnol. 62:61-68(2003).

DR EMBL; AY178757; AA031602.1; -

DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0006260; P:DNA replication; IEA.

DR InterPro; IPR000989; Rep.

DR Pfam; PF01446; Rep_1; 1.

KW Plasmid.

SQ SEQUENCE 379 AA; 40972 MW; AB53DB0C49736E0D CRC64;

Query Match 100.0%; Score 2005; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 1.8e-152;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTSVSAEHLGKDRPPVLVSDDKRGIRHELRPKLQITTTSETFNACGRPSGVNGVTIVN 60

Db 1 MTSVSAEHLGKDRPPVLVSDDKRGIRHELRPKLQITTTSETFNACGRPSGVNGVTIVN 60

Qy 61 GPKSGFGGLRSCGKGMICPCACGKGAHRADEISQVAHQLTGSGVAVYTMTRHTAGQ 120

Db 61 GPKSGFGGLRSCGKGMICPCACGKGAHRADEISQVAHQLTGSGVAVYTMTRHTAGQ 120

Qy 121 RLHDLWTGLSAAKAATNGRRWTEREMYGCDGYVRAVEITHGKNGHVVHALLMPSGD 180

Db 121 RLHDLWTGLSAAKAATNGRRWTEREMYGCDGYVRAVEITHGKNGHVVHALLMPSGD 180

Qy 181 VSENILESFSDAMFDRWTSKLSGLFAAPLRNSGGLDVRKIGGEADQVLAAYLTKIAGSV 240

Db 181 VSENILESFSDAMFDRWTSKLSGLFAAPLRNSGGLDVRKIGGEADQVLAAYLTKIAGSV 240

Qy 241 GMEVSGDGGKSGRHGNRAPEWIEAVDVGDPQALELWREFEFGSGMRRATANSRGLRARA 300

Db 241 GMEVSGDGGKSGRHGNRAPEWIEAVDVGDPQALELWREFEFGSGMRRATANSRGLRARA 300

QY	301	GLGAEITDAQIVQEESAPVVAIIIPARSWMIITCAPYVGEITGLVEAGATWENLRDH	360
Db	301	GLGAEITDAQIVQEESAPVVAIIIPARSWMIITCAPYVGEITGLVEAGATWENLRDH	360
QY	361	LHYRLPAADVRPPIISVRK	379
Db	361	LHYRLPAADVRPPIISVRK	379
RESULT 2			
P94171		PRELIMINARY;	PRT; 459 AA.
ID	P94171		
AC	P94171		
DT	01-MAY-1997	(Tremblrel. 03, Created)	
DT	01-MAY-1997	(Tremblrel. 03, Last sequence update)	
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)	
DE	Rep.		
GN	Name=rep;		
OS	Arcanobacterium pyogenes.		
OG	Plasmid pAPI.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Actinomycetaceae; Actinomycetaceae; Arcanobacterium.		
OX	NCBI_TaxID=1661;		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BBRI;		
RX	MEDLINE=98292760; PubMed=9620977;		
RA	Billington S.J., Jost B.H., Songer J.G.;		
RT	"The Arcanobacterium (Actinomycetes) pyogenes plasmid pAPI is a member		
RT	of the pU101/pJVI family of rolling circle replication plasmids."		
RL	J. Bacteriol. 180:3233-3236(1998).		
DR	EMBL; U83788; AAC6399.1; -.		
DR	GO; GO:0005727; C:extrachromosomal circular DNA; IEA.		
DR	GO; GO:0003677; F:DNA binding; IEA.		
DR	GO; GO:0006260; P:DNA replication; IEA.		
DR	InterPro; IPR000989; Rep.		
DR	Pfam; PF01446; Rep_1; 1.		
KW	Plasmid.		
SQ	SEQUENCE 459 AA; 51264 MW; 2A0A4B0A043BF865 CRC64;		
Query Match 30.4%; Score 609.5; DB 2; Length 459;			
Best Local Similarity 36.3%; Pred. No. 2.1e-40;			
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;			
QY	5	SAEHLSGKDRPPVLVSSD-----XRGIRHELRLPKLQIITSETFNACGR-PISGVN	56
Db	67	SATHPLGNTVLTFFPVSNESKTKAGRRSERYELRDGLAISTIESVRKCGRPVAPLVSL	126
QY	57	TVVNGPKSGFGGLRSCKGWLCPCCAGKVGGAHRADEISQVVAHOLGTG-SVAMVTMTMR	115
Db	127	RAKSGKGAGYGLHTCGSWACPVCSAKIAARRKTDLQVVVDHAVKHGNTVSMILTQ	186
QY	116	HTAGORLHDLTGLSAANKAAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL	175
Db	187	HHKGQGLKHLMDALSTAWNRVTSGRRIWIEFKQFGLGVYVRANEITHGKNGHVVSHVLI	246
QY	176	MFGSDVSENILSF-----SDAMFDRWTSKLVSGLGFAAPLRNSGG	215
Db	247	ISEKD---PLTSTFVYQKQRRRLPYPEIYMSDFIAERWEAGLAKHGVDFT-LRDSGG	302
QY	216	L-----DVRKTGGADQVLAAYLTKI---ASGVGMEVSGDGKSGRHNAPWEIAVDA	266
Db	303	LQWTAKDARATN-----YVSKQSTDAISSEVTLGGFKKARNGNRTPFQILLADI	354
QY	267	VG-GDPOALELWREFFSGMRRRAIAWSRGLRAGLGAELTDAQIVQEESAPVVAII	325
Db	355	LSIGVDVDDLKWKYEKASFGRRALTWSKGLRDWANLGVQSDDEEIA-SREIGDEAIALF	413
QY	326	PARSWMIITCAPYVGE--EILGLVEAGA-----TWNLRDLHLHYRLP	366
Db	414	THDAMQVRR-----FGAAELLDVTSGGRAAAYRWLDFRE-IDWSLP	455
RESULT 3			
Q8GDF1		PRELIMINARY;	PRT; 433 AA.
ID	Q8GDF1		
AC	Q8GDF1		
DT	01-MAR-2003	(Tremblrel. 23, Created)	
DT	01-MAR-2003	(Tremblrel. 23, Last sequence update)	
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)	
DE	Rep.		
GN	Name=rep;		
OS	Propionibacterium granulosum.		
OG	Plasmid cryptic plasmid pPG01.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.		
OX	NCBI_TaxID=33011;		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PF283;		
RA	Farrar M.D., Holland K.T.;		
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY150274; AAN78123.1; -.		
DR	GO; GO:0005727; C:extrachromosomal circular DNA; IEA.		
DR	GO; GO:0003677; F:DNA binding; IEA.		
DR	GO; GO:0006260; P:DNA replication; IEA.		
DR	InterPro; IPR000989; Rep.		
DR	Pfam; PF01446; Rep_1; 1.		
KW	Plasmid.		
SQ	SEQUENCE 433 AA; 48340 MW; 1A8C34C0FBF8678F CRC64;		
Query Match 28.4%; Score 569; DB 2; Length 433;			
Best Local Similarity 35.6%; Pred. NO. 3.5e-37;			
Matches 136; Conservative 63; Mismatches 145; Indels 38; Gaps 14;			
QY	16	PVLVSSDKRGIRHELRLPKLQ--QITSETFNACGR-PISGVN	72
Db	51	PCRKNRGRHRYENRDLRNPQVPLRVRKCGAVPV--QRIALMAGHGAGYAGLAT	108
QY	73	CGKWCICCCAGKVGGAHRADEISQVVAHOLGTG-SVAMVTMTMRHTAGORLHDLTGLSA	131
Db	109	CGSVWACPVCAAKISAHRRDELARVVQVAVGLGFKVSMILTQRRHAGQDLAELWASLOS	168
QY	132	AWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALLMFSGD---VSENI---	185
Db	169	GNNAVTSGRWQEFCAQLGVQGVKXAVETHGSHGHVHVHLVISKQDPTSVDTKIRHR	228
QY	186	-----LESFSDAMFDRWTSKLVSGL--FAAPLRNSGGLDVRKIGGEADQVLA	230
Db	229	RKQGRRTTYPPEVORPEDTAERWSRGLRKGVDVFLA---GSGGLDQWQADSGDEALG	285
QY	231	AYLTKIAS---GVGMEVSGDGKSGRHNAPWEIAVDAVG-GDPOALELWREFFSGM	286
Db	286	RIVAKMNSVDCGLANEATLGGFKKARNGNRTPFQILEDFLDTGSETDLRLWRTYVSASHG	345
QY	287	RRAIAWSRGLRAGLGAELTDAQIVQEESAPVVAIIIPARSWMIITCAPYVGEI-	344
Db	346	RKALWTSKGLRDWAGMESEDEQVAAQDCGEA-VALPFDHAWQIRTAGAFLDLELE	404
QY	345	LGLEVEAGATWENLRDLHLHYRLP	366
Db	405	LHGSEGVYAWLAKR-RIHYEIP	425
RESULT 4			
O69764		PRELIMINARY;	PRT; 403 AA.
ID	O69764		
AC	O69764		
DT	01-AUG-1998	(Tremblrel. 07, Created)	
DT	01-AUG-1998	(Tremblrel. 07, Last sequence update)	
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)	
DE	Putative rep protein.		
GN	Name=rep;		
OS	Propionibacterium freudenreichii.		
OG	Plasmid pLMB108.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		

OC	Propionibacterineae; Propionibacteriaceae; Propionibacterium.
OX	NCBI_taxID=1744;
FN	[1]
RP	SEQUENCE FROM N.A.
RA	Dasen G.H., Miescher S., Teuber M., Meile L.;
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AJ006662; CAA07175.1; --
DR	GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR	GO; GO:0003677; F:DNA binding; IEA.
DR	GO; GO:0006260; P:DNA replication; IEA.
DR	InterPro; IPR000989; Rep.
DR	Pfam; PF01446; Rep_1; 1.
KW	Plasmid.
QY	SEQUENCE 403 AA; 43502 MW; 340F40AA697E880C CRC64;
Query Match 24.8%; Score 497; DB 2; Length 403;	
Best Local Similarity 34.0%; Pred. No. 2e-31;	
Matches 127; Conservative 63; Mismatches 144; Indels 40; Gaps 15	
Qy	19 VSSDKRIRHRLRPKLOQITTSFTNACGR-PISGVNGVTIVNGPKSGFGGLRSCGKGW 77
Db	47 VAADKRHRFSVRYLWHRHTSLKRVAFGCRVAASAVASGVRCSDGRAGFAGLQCGSVW 106
Qy	78 ICPCACAKVGAHRADEISQVV----AHLGLTGSVAMVTWTRHTAGQRLDHLDTGLSAAW 133
Db	107 ACPVCNAKIWARGLGLEAAVETWTKH---GGRVAFMTFTVHRSKDLSLTAVWDGVASGW 163
Qy	134 KAATNGRRWRTEREMYGCDGVYRAVEITHGKNGWHVHVHALLMPSGDVSENLBSFSDAM 193
Db	164 RRVTSIGKGTSDQLRHGVEGEFVRVVEVTHGNGWHVHLVFLVLGDFGDAL--ALHRSM 221
Qy	194 FDRWTSKLVLGFAPIRNSGGLDVRIKIGBAD-OVLAAYLTK--IASGVGMEVSGDGK 250
Db	222 FGRWEGVLAAGLTGPTLARA--QDVQMSAATGLDHLARVLSKAQFGQKIGHELTNSQSK 279
Qy	251 SGRH--GNRAPWETAVDVGDDPALELWREFEFGSMGRRAIANSRGLRAPAGLGLTLD 308
Db	280 TARSALSTRTWELVDGAANGLAKEVGLWFEWKGSRGRQIGWSAGLRDLRLGLMWBESD 339
Qy	309 AQIVEQF-ESAPVNVVAIIPARSNW-----MIRTCAPYVFEILGLVEAG-ATWEN 356
Db	340 DVLAAEVGVSDVTVALITGDGWRIRVQOKLYECLRSC-----ELGG--QAGLARW-- 389
Qy	357 LRDH-LHYRLPAD 369
Db	390 LOEHGIEHELVEVD 403

```

RESULT 5
QB84GF4
ID QB84GF4 PRELIMINARY; PRT; 482 AA.
AC QB84GF4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE RepA.
GN Name=repA;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OG Plasmid pAG3.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
NCBI_TaxID=1718;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=22220;
RX MEDLINE=22830013; PubMed=12948627;
RT Tauch A., Puhler A., Kalinowski J., Thierbach G.;
RA "Plasmids in Corynebacterium glutamicum and their molecular
RL classification by comparative genomics.";
RT J. Biotechnol. 104:27-40(2003).
DR EMBL; AY172684; AAC18194.1; -.
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.

```

```

DR GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000989; Rep.
DR Pfam; PF01446; Rep_1; 1.
DR KW Plasmid.
SQ SEQUENCE      482 AA; 52145 MW;  F5D493A32F5C7F6C CRC64;

Query Match      21.2%; Score 425.5; DB 2; Length 482;
Best Local Similarity 29.1%; Pred. No. 1.3e-25;
Matches 109; Conservative 75; Mismatches 138; Indels 53; Gaps 14.

QY      2  TSVSAEHLGKDRPVLVSSDKRCIRHELHPKLOQITTSFTFNACGRPISGVNGV-TIVN 60
Db      31  TTDTTKHKVKALSPALF-----RAKLWDLQRTMKWKTEDKTLACGRWVAGGSVASILW 86
QY      61  GPKG-SGFGGLRSCGKIWCPCACGAKGVGAHRDEISQVV-----AHQLGTGS-- 105
Db      87  QAQGRARWGLQNSHWSVGSPVASSVTCRRRAVEVDAAIQAWANGAGLTFAHVRGVSTSP 146
QY     107  -----VAMVTMTMRHTAQGRLHDLMTGLSAAKAAATNGRRWR----- 143
Db     147  SDRKQRGVSNPVVERGISLMTLTLRHNSKQSLTEVWDAIAGCQWQVNTAAWRGGARTA 206
QY     144  TEREMYGCDYVRAVEITHGKNGWHVHALLMFSGDVSNILESFSDAMPDRWTSKLVS 203
Db     207  GDKSRYGIAHWYRAIEVTHGKNGWHVHLVFLFHDRVLSVDERDSLADRVDFRWAAKAVR 266
QY     204  LGFAAPLURNSGGLDVRKIGGEADQV--LAAYLTK-IASGYCMVEVGSD-CKSGRHGNRAP 259
Db     267  LGWRAPSRDR-GIDVVHVAASSDDAKSIGYTCCKMLSGIAAEFTTTGQVTKAAGDNRTPT 325
QY     260  WEIADVAVGDDPOALE--LWREFEFGSMGRRAIAWSRGLRARAGLGAELTDAQIVEQ--E 315
Db     326  FQILGD-LGKQYTKRDHALWLEWEKSGKRRQTGWSQGTQKDLGIN-ELSDQIDDSLGD 383
QY     316  ESAPVMVAIIAPRSW 330
Db     384  NDQSEVVAMVGRGAW 398

```

```

RESULT 6
Q9RIQO PRELIMINARY; PRT; 476 AA.
ID AC Q9RIQO;
DT 01-MAY-2000 (TrEMBLrel_13, Created)
DT 01-MAY-2000 (TrEMBLrel_13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel_26, Last annotation update)
DE Replication protein.
DE Name:rep;
DE GN Name:rep;
OS Streptomyces natalensis.
OS Streptomyces natalensis.
OC Bacteria;
OC Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=68242;
RN [1]
RP SEQUENCE FROM N.A.
RA Mendes M.V., Aparicio J.F., Martin J.F.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AJ243257; CAB62261.1; --
DR GO: GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0006260; P:DNA replication; IEA.
DR InterPro: IPR000989; Rep.
DR Dfam: PF01446; Rep_1; 1.
KW Plasmid.
KW SEQUENCE 476 AA; 51890 MW; DEE4C8B3CC2B5B95 CRC64;
SQ
Query Match 21.2%; Score 424.5; DB 2; Length 476;
Best Local Similarity 29.7%; Pred. No. 1.5e-25;
Matches 121; Conservative 53; Mismatches 148; Indels 85; Gaps 14;
QY 37 ITTSETFNACGRPIGVNCGVTIVNGPKSGGF-GGLRSCGKGMICPCCGAKVGAHRADEFS 95
DB 1 MTTDKALSCGRYALG-GGVTPKVPFGGTAYVLAGLATCGKWHICPCCGAKIRSARTVEIQ 59

```

```
QY 96 QV-----VAHOLGTGSGVAMVTWHT 117
Db 60 AAGTAWEDIGINGPKIPARAANKOINDRTABDWAABEADGLOAGGLMLTLWRY 119
QY 118 AGORLHDLTCLSAANKAATN--GRWRTEREMYGCDGYVRAVEITHGK-NGWHVHVA 173
Db 120 SRHTLAEIVTOORDAWKALCONAGRWKAKKIDYGVVGFVRAVEVYGEANGWHPWHV 179
QY 174 LLMFSGDVSENLBSFSDAMPDRWTSKLVSLGFAAPLENSG-GLDVRKIGGEADQVLAAY 232
Db 180 LVFEDKPLTPQGDALAEVYEANSTALQDVGVLPDREHGVRLDLSHGEGGPR--ARY 237
QY 233 LTKTASG-----VMEVSGSGDKSGRHNAPWEIA-----VDAVGGD-POALELWREFE 281
Db 238 LMKYQDGKAAWTAAEMTRTDTKAGRDGHRTPFEIARVLLTEDAADDDRAQTVRLWQYE 297
QY 282 FSGMGRRAIANSRGIRAEAGLGAEL---TDAQIVQEESAPVMVAII-----PA 327
Db 298 TAARGMRLYWSNGRLKLAALVELDTRTDGEIAAEERQGEALAVILADPHWHQIARRKG 357
QY 328 RSWMMIRTCAPYVFEILGLVEA-GATWENLRDLHLHYELPAADVRRPP 373
Db 358 RSLQLLKAERGGQKVRAVESWGLVW-----GRDVLPP 392

RESULT 7
Q74211 PRELIMINARY; PRT; 525 AA.
AC Q74211
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP0854;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017229; AAS03171.1; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 525 AA; 56735 MW; DB5E83E68D35FD4 CRC64;
```

```
Query Match 20.4%; Score 409; DB 2; Length 525;
Best Local Similarity 28.0%; Pred. No. 3e-24;
Matches 125; Conservative 58; Mismatches 136; Indels 128; Gaps 18;

QY 22 DKRGIRHELKPKLQITTSFENACGR-----PISGVNGVTI-----VNGPKSGSGFG 68
Db 62 ERRIRWGARMWLQWASSLKAVRCGRVLHNDVAGDPDG-OGVVKREVDRMVASLH 120
QY 69 GLRSCGKWCPCACAGKGAHRADEISQV--AHOLGTGSGVAMVTWHTAGQRIHDLW 126
Db 121 GLMTCGSWACPRCSAVIANTRAAGIATVRECVRG-GRVYLLTLTMRHSRRDGLADLW 179
QY 127 TGLSAANKAATNGRWRTER-----EMYGCDGYVRAVEITHGK-- 164
Db 180 DSLSTAWRSVFGTRNWTGKERMVQRRRGLALLPEIMGDAERFDIAGVTRVVEATYGPKE 239
QY 165 ---NGWHVHVALMFSGDVSENLBS-----FSDAMFRWTSKLVS 203
Db 240 LGHGHWLHIALVFSVTSLSGLLIEGIERTLGRGVNHDWLARNVFAARIHQWWSQGLAK 299
QY 204 LGFAAPLRNSGGLDVRKIGGEADQVLAAYTK-----IASGVMEVSGS---DGKSGRGG 255
Db 121 GLMTCGSWACPRCSAVIANTRAAGIATVRECVRG-GRVYLLTLTMRHSRRDGLADLW 179
QY 127 TGLSAANKAATNGRWRTER-----EMYGCDGYVRAVEITHGK-- 164
Db 180 DSLSTAWRSVFGTRNWTGKERMVQRRRGLALLPEIMGDAERFDIAGVTRVVEATYGPKE 239
QY 165 ---NGWHVHVALMFSGDVSENLBS-----FSDAMFRWTSKLVS 203
Db 240 LGHGHWLHIALVFSVTSLSGLLIEGIERTLGRGVNHDWLARNVFAARIHQWWSQGLAK 299
QY 204 LGFAAPLRNSGGLDVRKIGGEADQVLAAYTK-----IASGVMEVSGS---DGKSGRGG 255
Db 300 AGCOMP--GSVAVDVREIDDEGAEGYVGYLSKATYDVAARIGLEVAGVSTKDARAER-- 355
```

```
QY 256 NRAPWEI-----AVDAVG-----GD-----POALELW 277
Db 356 NQTFEVLANLAESVDARGGIRTPRHAVLPAGNGDWAVIDSDTGEVASITAPGQWKV 415
QY 278 REFEFGSMGRRAIANSR-----GLRARAGLGAELTDAQIVQEESAPVMVA 323
Db 416 HEWEQASCGRRQITWSRRRNPESGRBMLMNDLLDSRGSAAESDEIAVDEVDAB-SVG 474
QY 324 IIPARSWMMIRTCAPYVFEILGLVEA 350
Db 475 VISRQWYQVFAWRP---GLIVDLLEA 498

RESULT 8
AAS03171 PRELIMINARY; PRT; 525 AA.
AC AAS03171;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN MAP0854.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017229; AAS03171.1; -
KW Hypothetical protein.
SQ SEQUENCE 525 AA; 56735 MW; DB5E83E68D35FD4 CRC64;
```

```
Query Match 20.4%; Score 409; DB 2; Length 525;
Best Local Similarity 28.0%; Pred. No. 3e-24;
Matches 125; Conservative 58; Mismatches 136; Indels 128; Gaps 18;

QY 22 DKRGIRHELKPKLQITTSFENACGR-----PISGVNGVTI-----VNGPKSGSGFG 68
Db 62 ERRIRWGARMWLQWASSLKAVRCGRVLHNDVAGDPDG-OGVVKREVDRMVASLH 120
QY 69 GLRSCGKWCPCACAGKGAHRADEISQV--AHOLGTGSGVAMVTWHTAGQRIHDLW 126
Db 121 GLMTCGSWACPRCSAVIANTRAAGIATVRECVRG-GRVYLLTLTMRHSRRDGLADLW 179
QY 127 TGLSAANKAATNGRWRTER-----EMYGCDGYVRAVEITHGK-- 164
Db 180 DSLSTAWRSVFGTRNWTGKERMVQRRRGLALLPEIMGDAERFDIAGVTRVVEATYGPKE 239
QY 165 ---NGWHVHVALMFSGDVSENLBS-----FSDAMFRWTSKLVS 203
Db 240 LGHGHWLHIALVFSVTSLSGLLIEGIERTLGRGVNHDWLARNVFAARIHQWWSQGLAK 299
QY 204 LGFAAPLRNSGGLDVRKIGGEADQVLAAYTK-----IASGVMEVSGS---DGKSGRGG 255
Db 300 AGCOMP--GSVAVDVREIDDEGAEGYVGYLSKATYDVAARIGLEVAGVSTKDARAER-- 355
QY 256 NRAPWEI-----AVDAVG-----GD-----POALELW 277
Db 356 NQTFEVLANLAESVDARGGIRTPRHAVLPAGNGDWAVIDSDTGEVASITAPGQWKV 415
QY 278 REFEFGSMGRRAIANSR-----GLRARAGLGAELTDAQIVQEESAPVMVA 323
Db 416 HEWEQASCGRRQITWSRRRNPESGRBMLMNDLLDSRGSAAESDEIAVDEVDAB-SVG 474
QY 324 IIPARSWMMIRTCAPYVFEILGLVEA 350
Db 475 VISRQWYQVFAWRP---GLIVDLLEA 498
```

RESULT 9


```

Db          400 AN 401
|:
RESULT 14
Q9R881      PRELIMINARY;      PRT;      401 AA.
AC          Q9R881;
DT          01-MAY-2000 (TReMBLrel. 13, Created)
DT          01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT          01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE          Replicase.
GN          Name=repB11;
OS          Corynebacterium glutamicum (Brevibacterium flavum).
OG          Plasmid pB11.
OC          Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC          Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX          NCBI_TaxID=1718;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ATCC 21086;
RX          MEDLINE=96276208; PubMed=8693028;
RA          Ankri S., Reyes O., Leblon G.;
RT          "Electrotransformation of highly DNA-restrictive corynebacteria with
RT          synthetic DNA.";
RL          plasmid 35:62-66(1996).
RN          [2]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ATCC 21086;
RA          Favey S., Leblon G., Reyes O.;
RL          Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN          [3]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ATCC 21086;
RA          Ankri S., Reyes O., Leblon G.;
RL          Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR          EMBL; AF092037; AAD08690.1; -.
KW          Plasmid.
SQ          SEQUENCE      401 AA; 44481 MW; 5E0590D936132C4D CRC64;

Query Match      16.3%; Score 326; DB 2; Length 401;
Best Local Similarity 29.1%; Pred. No. 1e-17;
Matches 98; Conservative 58; Mismatches 131; Indels 50; Gaps 15;

Qy      34 LQQTITSETFNACGR-----PISGVNGVTIVNGPKSGFGLRSCGKWCIFCCAGKV 86
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MYKITNSKALAGCHRRRRDEAVVSSG-----NG--ASQPEGLQNGSHRWGSPLAIELE 53
Qy      87 GAHRADEISQVAHQLTGSGVAMVTM-TMRHTAGQRLHDLWTGLSAAKAATNGRRWRTE 145
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      54 MGERIELATATKNHLAAGNALMFMVGTVRNRSQSFQAQVAGIKTAYSSMVKTSQWKKE 113
Qy      146 REMYGCQGYVRVEITHG-KNGWHVHVHALLMFSGDVSENILESFSDAMFDRWTSKLVS 204
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      114 RARYGVHETYSDEYVETDSWANGWHLHRNMLLFLDRPLSDDELKAFEDSMFSRWGSAVVKA 173
Qy      205 GFAAPLNSG-GLD-VRKIGGEADQVLAAYLTKTASGVGMEVGGDGKSGRHGNRAPEWI 262
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      174 GMDAPLREHGVKLDQVSTWGGDAK-NATYLAKE---GMSQELTGSATKTASKGTYTPFQM 229
Qy      263 AVDAVGDDPOALE-----LWREFEFGSGMRAIAWSRGLRARAGIAELTDAQI--- 311
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      230 -LMDAQSDGENDDAVLVARWREYEVGSKNLRS-SWSRG--AKRALGIDYIDAVRRE 285
Qy      312 VEQE-----ESAPVMVAIIIPARSWMMIRT 335
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      286 MEELYKLGLAEAPERVESTREVALVKPDWKLIQS 322

RESULT 15
REP_STRLI
ID -REP STRLI
AC P22406; STANDARD; PRT; 456 AA.

```

```

DT          01-AUG-1991 (Rel. 19, Created)
DT          01-AUG-1991 (Rel. 19, Last sequence update)
DT          05-JUL-2004 (Rel. 44, Last annotation update)
DE          Rep protein.
GN          Name=rep;
OS          Streptomyces lividans.
OG          Plasmid pJ101.
OC          Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC          Streptomycineae; Streptomycetaceae; Streptomyces.
OX          NCBI_TaxID=1916;
RN          [1]
RP          SEQUENCE FROM N.A.
RX          MEDLINE=89008081; PubMed=3170481;
RA          Kendall K.J., Cohen S.N.;
RT          "Complete nucleotide sequence of the Streptomyces lividans plasmid
RT          pJ101 and correlation of the sequence with genetic properties.";
RL          J. Bacteriol. 170:4634-4651(1988).
CC          -!- FUNCTION: Essential for the autonomous replication of the plasmid
CC          pJ101.
CC          -----
CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          the European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC          or send an email to license@sib-sib.ch).
CC          -----
DR          EMBL; M21778; AAA88404.1; -.
DR          PIR; A31844; A31844.
DR          InterPro; IPR000989; Rep.
DR          Pfam; PF01446; Rep_1; 1.
DR          Plasmid; Plasmid partition.
KW          Plasmid; Plasmid partition.
SQ          SEQUENCE      456 AA; 49617 MW; 0976FD9D9429C7DC CRC64;

Query Match      15.8%; Score 316; DB 1; Length 456;
Best Local Similarity 30.5%; Pred. No. 7.5e-17;
Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

Qy      49 PISGV-----NGVTIVNGPKSGFGLRSCGKWCIFCCAGKVGHAHRADEISQVAHQ 102
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      3 PASGVIYAQTAAGTSVVL-----GLMRCGRWLCPVCAATIRHKRAEITAAVVEWI 54
Qy      103 GTGSVA-MVTWTRHTAGQRLHDLWTGLSAAKAATNGRR-----MRE----- 145
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      55 KRGGTAYLVFTARHGHGTRDLADLMDALQGTRKTPDSPRPGAYQRLITGCTWAGRAKD 114
Qy      146 -----REMYGCQGYVRVEITHG-KNGWHVHVHALLMFSGDV-----SENILESF 189
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      115 GHRAADREGIRDRIGYVGMIRATEVTVGQINGWHPHIAIVLGGRTGERSAKQIVATF 174
Qy      190 --SDAMFDRWTSKLVSGLGFAAPLR-----NSGGLDVRKIGGEAD-QVLAAYLTAKIA 237
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      175 EPTGAALDEWQGHWSV-WTAALRKVNPATPDDRHGVDPKRLETERDANDLAEYIAKTQ 233
Qy      238 SG--VGMVEVSGDGKSGRHGNRAPEWI---AVDAVG---DPQA-----IELWREFE 281
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      234 DGKAPALELARADIKTATGNNVAPFELLGRIGDLTGCTDADAAGVGSLEWNLRSRWHEYE 293
Qy      282 FGSNGRAIAWSRGLRARAGIAELTDA 309
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      294 RATRGRRAIEWTRYLQMLGLDGGDTEA 321

Search completed: October 23, 2004, 03:57:40
Job time : 112 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 22:14:09 ; Search time 5028 Seconds
(without alignments)
10722.008 Million cell updates/sec

Title: US-10-007-527A-1
Perfect score: 1140
Sequence: 1 atgaccagcgaagtgtgta.....taatatcggttcgaagtga 1140

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_btg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1140	100.0	1140	6	AX548642
C 2	1140	100.0	6334	1	AV178757
C 3	1140	100.0	6334	6	AX548646
C 4	1140	100.0	9652	6	AX548648
C 5	1140	100.0	11241	6	AX548647
C 6	1140	100.0	11241	12	AY180162
C 7	94.4	8.3	3540	1	AY150274
C 8	88.8	7.8	2439	1	APU83788
C 9	55.8	4.9	4603	1	AY172684
C 10	51.6	4.5	2297	6	E17316
C 11	51.6	4.5	5750	1	AF085719
C 12	49	4.3	125020	9	AF429315
C 13	47.2	4.1	2051	1	PF86662
C 14	46.4	4.1	125020	9	AF429315
C 15	44.4	3.9	2000	6	AX655393
C 16	42.4	3.7	9367	1	SN4243257
C 17	41.4	3.6	13651	1	AE011844
C 18	41.2	3.6	12855	1	AF533985
C 19	40.6	3.6	1189	6	AX826928

20	40.6	3.6	1244	6	AX826927	Sequence
21	40.6	3.6	1841	6	AX700509	Sequence
22	40.6	3.6	1841	10	MMPTX3	X83601 Mus musculus
C 23	40.6	3.6	163132	2	AC121312	AC121312 Mus muscu
C 24	39.2	3.4	127188	8	AC144481	AC144481 Medicago
C 25	38.8	3.4	248154	2	AC111496	AC111496 Rattus no
C 26	38.8	3.4	249495	2	AC095394	AC095394 Rattus no
C 27	38.6	3.4	125244	2	AC150093	AC150093 Gallus ga
C 28	38.6	3.4	230167	2	AC150046	AC150046 Gallus ga
C 29	38.6	3.4	275036	2	AC150059	AC150059 Gallus ga
C 30	38.4	3.4	205178	2	AC132020	AC132020 Rattus no
C 31	38.4	3.4	228187	2	AC132020	AC132020 Rattus no
C 32	38.4	3.4	304078	2	AC129703	AC129703 Rattus no
C 33	38.2	3.4	6758	1	AY172685	AY172685 Corynebac
C 34	37.6	3.3	122934	9	AC092898	AC092898 Homo sapi
C 35	37.4	3.3	2708	6	AX700512	AX700512 Sequence
C 36	37.4	3.3	2708	10	MMU33842	U33842 Mus musculu
C 37	37.4	3.3	3147	12	AF092035	AF092035 Integrati
C 38	37.4	3.3	4447	1	AF092037	AF092037 Corynebac
C 39	37.4	3.3	4447	6	AR173114	AR173114 Sequence
C 40	37.4	3.3	4447	6	AR173115	AR173115 Sequence
C 41	37.4	3.3	4447	6	BD106940	BD106940 Method fo
C 42	37.4	3.3	4447	6	BD106941	BD106941 Method fo
C 43	37.4	3.3	4447	6	BD139585	BD139585 Arginine
C 44	37.4	3.3	4447	6	BD139586	BD139586 Arginine
C 45	37.4	3.3	4457	1	BLPLASMI	X03987 Brevibacter

ALIGNMENTS

RESULT 1
AX548642
LOCUS AX548642 1140 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 1 from Patent WO02055709.
ACCESSION AX548642
VERSION AX548642.1 GI:25813612
KEYWORDS Rhodococcus erythropolis
SOURCE Rhodococcus erythropolis
ORGANISM Rhodococcus erythropolis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Nocardiaceae; Rhodococcus.
REFERENCE 1
AUTHORS Bramucci,M.G., Cheng,Q., Kostichka,K.N. and Tomb,J.F.
TITLE Rhodococcus cloning and expression vectors
JOURNAL Patent: WO 02055709-A 1 18-JUL-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1. .1140
/organism="Rhodococcus erythropolis"
/mol_type="unassigned DNA"
/strain="AN12"
/db_xref="taxon:1833"

ORIGIN

Query Match 100.0%; Score 1140; DB 6; Length 1140;
Best Local Similarity 100.0%; Pred. No. 1.1e-285;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACGAGCGTAGTGTGCTGACACCTTCGGGAAAGACGGGCTCCGTCCTCGTGTGCG 60
Db 1 ATGACGAGCGTAGTGTGCTGACACCTTCGGGAAAGACGGGCTCCGTCCTCGTGTGCG 60
QY 61 TCCGATAAGCGGGCATCCGGACGAACTGCGACCCAACTTCAACAAATCACCAGTCA 120
Db 61 TCCGATAAGCGGGCATCCGGACGAACTGCGACCCAACTTCAACAAATCACCAGTCA 120
QY 121 GAAACATTAAACGCTGTGTGCGCGCGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 180
Db 121 GAAACATTAAACGCTGTGTGCGCGCGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 180
QY 181 GGTCCGAAAGGTTCTGGATTCGAGGCGCTTGGTCTCTCGGAAAGGGGTGGATCTCGCCC 240

QY 181 GGTCCGAAAGGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 240
Db |||||
QY 2871 GGTCCGAAAGGTTCTGGATTCCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 2812
Db |||||
QY 241 TGCTGTGCGGAAAAAGTCCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db |||||
QY 2811 TGCTGTGCGGAAAAAGTCCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 2752
Db |||||
QY 301 CAATCCGGAGCTGGATCTGTTGGCATGTGGAGTGCAGATGACATGCGGCATACAGCTGGTCA 360
Db |||||
QY 2751 CAATCCGGAGCTGGATCTGTTGGCATGTGGAGTGCAGATGACATGCGGCATACAGCTGGTCA 2692
Db |||||
QY 361 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGCTGCGACCAACGCTCGT 420
Db |||||
QY 2691 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGCTGCGACCAACGCTCGT 2632
Db |||||
QY 421 GGTGGGCTACGGAACGTTGAAATGTACGGCTGCGACGGATACGTCGCGCTGTTGAAATC 480
Db |||||
QY 2631 CGTTGGGCTACGGAACGTTGAAATGTACGGCTGCGACGGATACGTCGCGCTGTTGAAATC 2572
Db |||||
QY 481 ACTCACGGAAAAACGGCTGGCAGCTGCACGTTTCCAGCGCTACTCATGTTCAAGTGGTCA 540
Db |||||
QY 2571 ACTCACGGAAAAACGGCTGGCAGCTGCACGTTTCCAGCGCTACTCATGTTCAAGTGGTCA 2512
Db |||||
QY 541 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGATGTTTCGATCGGCTGACTTCCAAA 600
Db |||||
QY 2511 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGGATGTTTCGATCGGCTGACTTCCAAA 2452
Db |||||
QY 601 CTCGTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 660
Db |||||
QY 2451 CTCGTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 2392
Db |||||
QY 661 ATCGCGGCTGAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTCGATCTGGGCTT 720
Db |||||
QY 2391 ATCGCGGCTGAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTCGATCTGGGCTT 2332
Db |||||
QY 721 GGTATGAGGTTGTTAGTGGGACGCGAAAAAGTGTGACATGCGACCGTGCACCCCTGG 780
Db |||||
QY 2331 GGTATGAGGTTGTTAGTGGGACGCGAAAAAGTGTGACATGCGACCGTGCACCCCTGG 2272
Db |||||
QY 781 GAAATCGCTGTTGATGAGTGGGCGGGATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 840
Db |||||
QY 2271 GAAATCGCTGTTGATGAGTGGGCGGGATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 2212
Db |||||
QY 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCGTTGATGCGTGGCCGAGCT 900
Db |||||
QY 2211 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCGTTGATGCGTGGCCGAGCT 2152
Db |||||
QY 901 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAAGAACTCTGCCCGGTC 960
Db |||||
QY 2151 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAAGAACTCTGCCCGGTC 2092
Db |||||
QY 961 ATGTTTGGATCAATTCGGGCGGATCGTGGATGATGATTCGGAATTCGTCGCGCTTACGTC 1020
Db |||||
QY 2091 ATGTTTGGATCAATTCGGGCGGATCGTGGATGATGATTCGGAATTCGTCGCGCTTACGTC 2032
Db |||||
QY 1021 TTGGGCGAGATCTCGGACTCGTGCAGCTGGCGGCTGCGGAAATTCGTTGATGATC 1080
Db |||||
QY 2031 TTGGGCGAGATCTCGGACTCGTGCAGCTGGCGGCTTGGGAAATTCGTTGATGATC 1972
Db |||||
QY 1081 TTGCATTATCATTTGCCGAGCGGATGTGCGGCCCGGATAATATCGTTTCCGCAAGTGA 1140
Db |||||
QY 1971 TTGCATTATCATTTGCCGAGCGGATGTGCGGCCCGGATAATATCGTTTCCGCAAGTGA 1912
Db |||||

RESULT 3
AX548646/c
LOCUS AX548646 6334 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 5 from Patent WO02055709.
ACCESSION AX548646
VERSION AX548646.1 GI:25813614
KEYWORDS Rhodococcus erythropolis
SOURCE

ORGANISM Rhodococcus erythropolis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Nocardiaceae; Rhodococcus.
REFERENCE
AUTHORS Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.
TITLE Rhodococcus cloning and expression vectors
JOURNAL Patent: WO 02055709-A 5 18-JUL-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1. 6334
/organism="Rhodococcus erythropolis"
/mol_type="unassigned DNA"
/strain="AN12"
/db_xref="taxon:1833"

ORIGIN

Query Match 100.0%; Score 1140; DB 6; Length 6334;
Best Local Similarity 100.0%; Pred. No. 1e-285;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACACGCGTAACTGCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
Db |||||
QY 3051 ATGACACGCGTAACTGCTGAAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTGCG 2992
Db |||||
QY 61 TCCGATAAGCGCGCATCCGGCAGCAAGCTGCGACCCAACTTCAACAAATCACCACGTCA 120
Db |||||
QY 2991 TCCGATAAGCGCGCATCCGGCAGCAAGCTGCGACCCAACTTCAACAAATCACCACGTCA 2932
Db |||||
QY 121 GAAACATTTAAAGCGCTGTCGGCGGCGATTTCTGCGTGAACGGTGTGACATTTGTCAC 180
Db |||||
QY 2931 GAAACATTTAAAGCGCTGTCGGCGGCGATTTCTGCGTGAACGGTGTGACATTTGTCAC 2872
Db |||||
QY 181 GGTCCGAAAGGTTCTGGATTTCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 240
Db |||||
QY 2871 GGTCCGAAAGGTTCTGGATTTCGAGGCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 2812
Db |||||
QY 241 TGCTGTGCGGAAAAAGTCCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db |||||
QY 2811 TGCTGTGCGGAAAAAGTCCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT 2752
Db |||||
QY 301 CAATCCGGAGCTGGATCTGTTGGCATGTGGAGTGCAGATGACATGCGGCATACAGCTGGTCA 360
Db |||||
QY 2751 CAATCCGGAGCTGGATCTGTTGGCATGTGGAGTGCAGATGACATGCGGCATACAGCTGGTCA 2692
Db |||||
QY 361 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGCTGCGACCAACGCTCGT 420
Db |||||
QY 2691 CGGCTCCACGACCTATGAGCTGGACTTTCGGCAGCTCGGAAAGCTGCGACCAACGCTCGT 2632
Db |||||
QY 421 GGTGGGCTACGGAACGTTGAAATGTAGCGCTGCGACGGATACGTCGCGCTGTTGAAATC 480
Db |||||
QY 2631 GGTGGGCTACGGAACGTTGAAATGTAGCGCTGCGACGGATACGTCGCGCTGTTGAAATC 2572
Db |||||
QY 481 ACTCACGGAAAAACGGCTGGCAGCTGCACGTTTCCAGCGCTACTCATGTTCAAGTGGTCA 540
Db |||||
QY 2571 ACTCACGGAAAAACGGCTGGCAGCTGCACGTTTCCAGCGCTACTCATGTTCAAGTGGTCA 2512
Db |||||
QY 541 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGATGTTTCGATCGGCTGACTTCCAAA 600
Db |||||
QY 2511 GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGGATGTTTCGATCGGCTGACTTCCAAA 2452
Db |||||
QY 601 CTGATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 660
Db |||||
QY 2451 CTGATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 2392
Db |||||
QY 661 ATCGCGGCTGAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTCGATCTGGGCTT 720
Db |||||
QY 2391 ATCGCGGCTGAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTCGATCTGGGCTT 2332
Db |||||
QY 721 GGTATGAGGTTGTTAGTGGGACGCGAAAAAGTGTGACATGCGACCGTGCACCCCTGG 780
Db |||||
QY 2331 GGTATGAGGTTGTTAGTGGGACGCGAAAAAGTGTGACATGCGACCGTGCACCCCTGG 2272
Db |||||
QY 781 GAAATCGCTGTTGATGAGTGGGCGGGATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 840
Db |||||
QY 2271 GAAATCGCTGTTGATGAGTGGGCGGGATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 2212
Db |||||
QY 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCGTTGATGCGTGGCCGAGCT 900
Db |||||
QY 2211 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCGTTGATGCGTGGCCGAGCT 2152
Db |||||
QY 901 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAAGAACTCTGCCCGGTC 960
Db |||||
QY 2151 GGTCTTGGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAAGAACTCTGCCCGGTC 2092
Db |||||
QY 961 ATGTTTGGATCAATTCGGGCGGATCGTGGATGATGATTCGGAATTCGTCGCGCTTACGTC 1020
Db |||||
QY 2091 ATGTTTGGATCAATTCGGGCGGATCGTGGATGATGATTCGGAATTCGTCGCGCTTACGTC 2032
Db |||||
QY 1021 TTGGGCGAGATCTCGGACTCGTGCAGCTGGCGGCTGCGGAAATTCGTTGATGATC 1080
Db |||||
QY 2031 TTGGGCGAGATCTCGGACTCGTGCAGCTGGCGGCTTGGGAAATTCGTTGATGATC 1972
Db |||||
QY 1081 TTGCATTATCATTTGCCGAGCGGATGTGCGGCCCGGATAATATCGTTTCCGCAAGTGA 1140
Db |||||
QY 1971 TTGCATTATCATTTGCCGAGCGGATGTGCGGCCCGGATAATATCGTTTCCGCAAGTGA 2332
Db |||||
QY 721 GGTATGAGGTTGTTAGTGGGACGCGAAAAAGTGTGACATGCGACCGTGCACCCCTGG 780
Db |||||
QY 2331 GGTATGAGGTTGTTAGTGGGACGCGAAAAAGTGTGACATGCGACCGTGCACCCCTGG 2272
Db |||||
QY 781 GAAATCGCTGTTGATGAGTGGGCGGGATCCACAAGCGTTTGGAACTGTGCGCAGAAATTT 840
Db |||||

Db 2271 GAAATCGTGTGATGAGTGGGCGGGATCCACAAGCTTGGAACTTGTGCGGAGATTT 2212
Qy 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGAATGCGTGGCCGAGCT 900
Db 2211 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGAATGCGTGGCCGAGCT 2152
Qy 901 GGTCTTGGGGGAGAACTAACAGATGCTCAGATGCTGATGAGTGGAGGAGAAATCTGCCCGGTC 960
Db 2151 GGTCTTGGGGGAGAACTAACAGATGCTCAGATGCTGATGAGTGGAGGAGAAATCTGCCCGGTC 2092
Qy 961 ATGGTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTTGGCGCCCTTACGTC 1020
Db 2091 ATGGTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTTGGCGCCCTTACGTC 2032
Qy 1021 TTGGCGGAGATCTCGGACTCGTGAAGCTCGCGGACTTGGGAAAATCTTCGTGATCAC 1080
Db 2031 TTGGCGGAGATCTCGGACTCGTGAAGCTCGCGGACTTGGGAAAATCTTCGTGATCAC 1972
Qy 1081 TTGCATTATCGATTGCCCGGAGCGATGTGGCGCCCGATTAATATCGTTTGGCAAGTGA 1140
Db 1971 TTGCATTATCGATTGCCCGGAGCGATGTGGCGCCCGATTAATATCGTTTGGCAAGTGA 1912

RESULT 4
AX548648/c 9652 bp DNA linear PAT 27-NOV-2002
LOCUS AX548648
DEFINITION Sequence 7 from Patent WO02055709.
ACCESSION AX548648
VERSION AX548648.1 GI:25813616
KEYWORDS
SOURCE Shuttle vector pRHB17
ORGANISM Shuttle vector pRHB17
artificial sequences; vectors.

REFERENCE
1
AUTHORS Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.
TITLE Rhodococcus cloning and expression vectors
JOURNAL Patent: WO 02055709-A 7 18-JUL-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1..9652
/organism="Shuttle vector pRHB17"
/mol_type="unassigned DNA"
/db_xref="taxon:214938"

ORIGIN
Query Match 100.0%; Score 1140; DB 6; Length 9652;
Best Local Similarity 100.0%; Pred. No. 1e-285;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACGAGGTAAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTG 60
Db 5052 ATGACGAGGTAAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCCGTCTCGTGTG 4993
Qy 61 TCCGATAAGCGGGATCCGGACCACTTCCGACCACTTCAACAAATACCAAGTCA 120
Db 4992 TCCGATAAGCGGGATCCGGACCACTTCCGACCACTTCAACAAATACCAAGTCA 4933
Qy 121 GAAACATTAAACGCTGTGGCGGCGGATTTCTGGCGTGAACGGTGTGACCAATTGTCAAC 180
Db 4932 GAAACATTAAACGCTGTGGCGGCGGATTTCTGGCGTGAACGGTGTGACCAATTGTCAAC 4873
Qy 181 GGTCCGAAGTTCGATTCGGAGCGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 240
Db 4872 GGTCCGAAGTTCGATTCGGAGCGCTTCGTTCTCGGAAAGGCTGGATCTGCCCC 4813
Qy 241 TGCTGTGGGGGAAAAGTCGGTGCATCGTCGAGACGAAATTTCAAGTTCTTGCATCAT 300
Db 4812 TGCTGTGGGGGAAAAGTCGGTGCATCGTCGAGACGAAATTTCAAGTTCTTGCATCAT 4753
Qy 301 CAACTCGGAGCTGGATCTGTTGGATGGTGAAGTGAACCATGCGGCATACAGCTGGTCA 360
Db 4752 CAACTCGGAGCTGGATCTGTTGGATGGTGAAGTGAACCATGCGGCATACAGCTGGTCA 4693

Qy 361 CGGCTCCACGACCTATGACTGGACTTTTCGGACGCTTGGAAAGCTGGCAACCAACGGTGGT 420
Db 4692 CGGCTCCACGACCTATGACTGGACTTTTCGGACGCTTGGAAAGCTGGCAACCAACGGTGGT 4633
Qy 421 CGTTGGCGGTACGGAAACGTGAAATGATACGGCTTCGACGATACGTCGCGCTGTTGAAATC 480
Db 4632 CGTTGGCGGTACGGAAACGTGAAATGATACGGCTTCGACGATACGTCGCGCTGTTGAAATC 4573
Qy 481 ACTCACGGAAAAAAGCGCTGCGACGTCGATCGATCGTTCACGGCTTACTCATGTTTCAGTGGT 540
Db 4572 ACTCACGGAAAAAAGCGCTGCGACGTCGATCGATCGTTCACGGCTTACTCATGTTTCAGTGGT 4513
Qy 541 GTGAGTGCAGAACATCTCCGATCTCTTCGATCGGATGTTTCGATCGGATGACTTCCAAA 600
Db 4512 GTGAGTGCAGAACATCTCCGATCTCTTCGATCGGATGTTTCGATCGGATGACTTCCAAA 4453
Qy 601 CTCGTATCTCTGGGATTTGCTGCGCACTACGTAATTCGGGTGTTCTCGATGATGAGAAAG 660
Db 4452 CTCGTATCTCTGGGATTTGCTGCGCACTACGTAATTCGGGTGTTCTCGATGATGAGAAAG 4393
Qy 661 ATCGCGCGTGAAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTCGATCTGGCGCT 720
Db 4392 ATCGCGCGTGAAGCTGATCAAGTTCTCGCTCGGATCTGACGAAATTCGATCTGGCGCT 4333
Qy 721 GGTATGAGGTTTGGTGTGCGACGAAAGTGGTTCGACATGGCAACCGTGCAACCTGG 780
Db 4332 GGTATGAGGTTTGGTGTGCGACGAAAGTGGTTCGACATGGCAACCGTGCAACCTGG 4273
Qy 781 GAAATCGCTGTTGATGACGTGGCGGGGATCCAAAGCTTGGAACTCTGGCGAGAAATTT 840
Db 4272 GAAATCGCTGTTGATGACGTGGCGGGGATCCAAAGCTTGGAACTCTGGCGAGAAATTT 4213
Qy 841 GAGTTTGGTTCGATGGGACGTCGGCAATCGCGTGGTCCCGTGGATTCGGTCCCGAGCT 900
Db 4212 GAGTTTGGTTCGATGGGACGTCGGCAATCGCGTGGTCCCGTGGATTCGGTCCCGAGCT 4153
Qy 901 GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 4152 GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 4093
Qy 961 ATGGTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTTGGCGCTTACGTC 1020
Db 4092 ATGGTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTTGGCGCTTACGTC 4033
Qy 1021 TTGGCGGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAATCTTCGATGATCAC 1080
Db 4032 TTGGCGGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAATCTTCGATGATCAC 3973
Qy 1081 TTGCATTATCGATTGCCCGGAGCGATGTGGCGCCCGGATTAATATCGTTTCCGAAGTGA 1140
Db 3972 TTGCATTATCGATTGCCCGGAGCGATGTGGCGCCCGGATTAATATCGTTTCCGAAGTGA 3913

RESULT 5
AX548647/c 11241 bp DNA linear PAT 27-NOV-2002
LOCUS AX548647
DEFINITION Sequence 6 from Patent WO02055709.
ACCESSION AX548647
VERSION AX548647.1 GI:25813615
KEYWORDS
SOURCE Shuttle vector pRHB17
ORGANISM Shuttle vector pRHB17
artificial sequences; vectors.

REFERENCE
1
AUTHORS Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.
TITLE Rhodococcus cloning and expression vectors
JOURNAL Patent: WO 02055709-A 6 18-JUL-2002;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1..11241
/organism="Shuttle vector pRHB17"
/mol_type="unassigned DNA"

Db	5681	ATGTTTGCATCATTCGCGCGCATCTGTCGATGATGATTCGGACTTCGGACTTCGGCTTCACGTC	562
Qy	1021	TTGCGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTGGGAAAATCTTCGTGATCAC	108
Db	5621	TTGCGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTGGGAAAATCTTCGTGATCAC	5562
Qy	1081	TTGCATTATCATTTGCCCGCAGCGGATGTGCGGCCCGGATTAATATATCGTTTCGCAAGTGA	1140
Db	5561	TTGCATTATCATTTGCCCGCAGCGGATGTGCGGCCCGGATTAATATATCGTTTCGCAAGTGA	5502
RESULT 6			
AY180162/c		11241 bp DNA circular SYN 23-JUL-2003	
LOCUS		Shuttle vector pRHR17, complete sequence.	
DEFINITION		Shuttle vector pRHR17, complete sequence.	
ACCESSION		AY180162	
VERSION		AY180162.1 GI:30313707	
KEYWORDS		Shuttle vector pRHR17	
SOURCE		Shuttle vector pRHR17	
ORGANISM		artificial sequences; vectors.	
REFERENCE		1 (bases 1 to 11241)	
AUTHORS		Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and Cheng, Q.	
TITLE		A small cryptic plasmid from Rhodococcus erythropolis: characterization and utility for gene expression	
JOURNAL		Appl. Microbiol. Biotechnol. 62 (1), 61-68 (2003)	
MEDLINE		22718480	
PUBMED		12835922	
REFERENCE		2 (bases 1 to 11241)	
AUTHORS		Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and Cheng, Q.	
TITLE		Direct Submission	
JOURNAL		Submitted (14-NOV-2002) CR@D, E. I. Dupont de Numours Inc., Experimental Station, Wilmington, DE 19880-0328, USA	
FEATURES		Location/Qualifiers	
source		1. 11241	
		/organism="Shuttle vector pRHR17"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:214938"	
		/focus	
		/note="Rhodococcus - Escherichia coli shuttle vector; Ssp I linearize pAN12 cloned into the PvuII site of pBR328" join(9925..11241,1..3590)	
source		/organism="Cloning vector pBR328"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:221946"	
source		3591..9924	
		/organism="Rhodococcus erythropolis"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:1833"	
		/plasmid="pAN12"	
		/note="cryptic plasmid"	
ORIGIN			
Query Match		100.0%; Score 1140; DB 12; Length 11241;	
Best Local Similarity		100.0%; Pred. No. 1e-285;	
Matches 1140; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	ATGACGAGCTAAGTCTGAACACCTTTCGGCAAGACCGGCTCCCGTCTCGTGTGG	60
Db	6641	ATGACGAGCTAAGTCTGAACACCTTTCGGCAAGACCGGCTCCCGTCTCGTGTGG	6582
Qy	61	TCGGATAAGCGCGGCATCCGGCAGCACTGGACCCCAAACTTCAACAAATCACACGTCA	120
Db	6581	TCGGATAAGCGCGGCATCCGGCAGCACTGGACCCCAAACTTCAACAAATCACACGTCA	6522
Qy	121	GAAACATTTAACCGCTGTGGCGCGCGATTTCTGGCGTGAACCGGTGTGACCATTTGTCAC	180
Db	6521	GAAACATTTAACCGCTGTGGCGCGCGATTTCTGGCGTGAACCGGTGTGACCATTTGTCAC	6462
Qy	181	GGTCCGAAGGTTCTGGATTCGGAGGCTTCTGCTCGGAAAGGGCTGATCTGCCCC	240
Db	6461	GGTCCGAAGGTTCTGGATTCGGAGGCTTCTGCTCGGAAAGGGCTGATCTGCCCC	6402
Qy	241	TGCTGTGGGAAAGTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT	300
Db	6401	TGCTGTGGGAAAGTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT	6342
Qy	301	CAACTCGGACTGATCTGTTGCCATGTGACGATGACCATGCGGCATACAGCTGGTCAG	360
Db	6341	CAACTCGGACTGATCTGTTGCCATGTGACGATGACCATGCGGCATACAGCTGGTCAG	6282
Qy	361	CGGCTCCACGACCTATGACTGGATTTCCGACGCTTGGAAAGCTGGACCAACGGTCTGT	420
Db	6281	CGGCTCCACGACCTATGACTGGATTTCCGACGCTTGGAAAGCTGGACCAACGGTCTGT	6222
Qy	421	CGTTGGCGTACGGAACGTGAATGTACGGCTGCGACGCGATACGTGCGCGCTGTTGAAATC	480
Db	6221	CGTTGGCGTACGGAACGTGAATGTACGGCTGCGACGCGATACGTGCGCGCTGTTGAAATC	6162
Qy	481	ACTCAGCGAAACACGGCTGGCGACGTCCACGTTTCACCGCGTACTCATGTTCAAGTGGTAC	540
Db	6161	ACTCAGCGAAACACGGCTGGCGACGTCCACGTTTCACCGCGTACTCATGTTCAAGTGGTAC	6102
Qy	541	GTGAGTCAGAACATCTCTCGATCTCTCGATCGGATGTCGATCGGTGACATTCACAA	600
Db	6101	GTGAGTCAGAACATCTCTCGATCTCTCGATCGGATGTCGATCGGTGACATTCACAA	6042
Qy	601	CTCGTATCTCTGGGATTTGCTGGCCACTAGTAAATTCGGGTGCTCTCGATGACGAAAG	660
Db	6041	CTCGTATCTCTGGGATTTGCTGGCCACTAGTAAATTCGGGTGCTCTCGATGACGAAAG	5982
Qy	661	ATCGCGGTGAAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAATTCGATCTGGCGTT	720
Db	5981	ATCGCGGTGAAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAATTCGATCTGGCGTT	5922
Qy	721	GGTATGAGGTTTGGTATGCGGACGGAAGTGGTGCATGCAACCGTGCACCCCTGG	780
Db	5921	GGTATGAGGTTTGGTATGCGGACGGAAGTGGTGCATGCAACCGTGCACCCCTGG	5862
Qy	781	GAAATCGCTGTGATGAGTGGGCGGGATCCCAAGCGTTTGGAACTGTGGCGGAGATTT	840
Db	5861	GAAATCGCTGTGATGAGTGGGCGGGATCCCAAGCGTTTGGAACTGTGGCGGAGATTT	5802
Qy	841	GAGTTTGGTTCGATGGGACGTGGGGCAATCCGCTGGTCCCGTGTGATTCGTCGCCGAGCT	900
Db	5801	GAGTTTGGTTCGATGGGACGTGGGGCAATCCGCTGGTCCCGTGTGATTCGTCGCCGAGCT	5742
Qy	901	GCTCTGGGCGAGAACTAAACAGATGCTCAGATCGTTTCAGCAGGAGAAATCTGCCCGGTC	960
Db	5741	GCTCTGGGCGAGAACTAAACAGATGCTCAGATCGTTTCAGCAGGAGAAATCTGCCCGGTC	5682
Qy	961	ATGGTTCCGATCATTCGGGCGCATTCGTGGATGATGATTCGCACTTGTGCGCCTTACGTC	1020

Db	6461	GGTCCGAAAGTTCTCGATTCGAGGCTTCTGGTTCCTCGCGAAAGGCTGGATCTGCCCC	6402		
Qy	241	TGCTGTGCGGAAAAGTCGGTGACATCGTGCAGACGAAATTTCTCAAGTTGTGCTCAT	300		
Db	6401	TGCTGTGCGGAAAAGTCGGTGACATCGTGCAGACGAAATTTCTCAAGTTGTGCTCAT	6342		
Qy	301	CAACTCGGACTGATCTGTTGCGATGGTGACGATGACCAATGCGCCATACAGCTGGTCAG	360		
Db	6341	CAACTCGGACTGATCTGTTGCGATGGTGACGATGACCAATGCGCCATACAGCTGGTCAG	6282		
Qy	361	CGGCTCCACGACCTATGGAATGGAATTTTCGGCAGCCTGGAAGCTGCGACCAACGGTCTG	420		
Db	6281	CGGCTCCACGACCTATGGAATGGAATTTTCGGCAGCCTGGAAGCTGCGACCAACGGTCTG	6222		
Qy	421	CGTTGGCGTACGGAACGTTGAAATGTCAGGCTGCGACGGAATACGTGCGCGCTGTTGAAATC	480		
Db	6221	CGTTGGCGTACGGAACGTTGAAATGTCAGGCTGCGACGGAATACGTGCGCGCTGTTGAAATC	6162		
Qy	481	ACTCACGGAAGAAACGGCTGCGACGTCACGTTTCAACGCGTACTCATGTTTCAGTGGTGAC	540		
Db	6161	ACTCACGGAAGAAACGGCTGCGACGTCACGTTTCAACGCGTACTCATGTTTCAGTGGTGAC	6102		
Qy	541	GTGAGTGAGAACATCTCTCGAATCTCTCGATGCGATGTTTCGATCGGTGGACTTCCAAA	600		
Db	6101	GTGAGTGAGAACATCTCTCGAATCTCTCGATGCGATGTTTCGATCGGTGGACTTCCAAA	6042		
Qy	601	CTCGTATCTCTGGGATTTGCTGCGCCACTAGCTAATTCGGGTGCTCTCGATGACGAAG	660		
Db	6041	CTCGTATCTCTGGGATTTGCTGCGCCACTAGCTAATTCGGGTGCTCTCGATGACGAAG	5982		
Qy	661	ATCGCGCGTGAAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAATTCGATCTGGCGTT	720		
Db	5981	ATCGCGCGTGAAGCTGATCAAGTTCTCGCTGCGTATCTGACGAAATTCGATCTGGCGTT	5922		
Qy	721	GGTATGAGGTTTGTAGTGGCGACGGAAGAGTGTGCAATGCAACCGTGACCCCTGG	780		
Db	5921	GGTATGAGGTTTGTAGTGGCGACGGAAGAGTGTGCAATGCAACCGTGACCCCTGG	5862		
Qy	781	GAAATCGCTGTTGATGCGATGCGGCGGATCCCAAGCGTTTGGAACTGTGGCGAATTT	840		
Db	5861	GAAATCGCTGTTGATGCGATGCGGCGGATCCCAAGCGTTTGGAACTGTGGCGAATTT	5802		
Qy	841	GAGTTTGTGTCGATGGGACGTCGGGCAATCGCGTGTCCCGTGGATTCGTCGCCGAGCT	900		
Db	5801	GAGTTTGTGTCGATGGGACGTCGGGCAATCGCGTGTCCCGTGGATTCGTCGCCGAGCT	5742		
Qy	901	GCTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGATCTGCCCGGTC	960		
Db	5741	GCTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGATCTGCCCGGTC	5682		
Qy	961	ATGTTTGGCATCATTCGCGCGGATCGTGGATGATGATTCGGAATTTGCGGCTTACGTC	1020		
Db	5681	ATGTTTGGCATCATTCGCGCGGATCGTGGATGATGATTCGGAATTTGCGGCTTACGTC	5622		
Qy	1021	TTCGCGGATGCTCTCGACTCTCGAAGCTGCGCGCACTTGGGAAATCTTCGTGATCAC	1080		
Db	5621	TTCGCGGATGCTCTCGACTCTCGAAGCTGCGCGCACTTGGGAAATCTTCGTGATCAC	5562		
Qy	1081	TTGCATTATCGATTGCCCGCAGCGGATGTGCGGCCCCCGGATAATATCGGTTTCGCAAGTGA	1140		
Db	5561	TTGCATTATCGATTGCCCGCAGCGGATGTGCGGCCCCCGGATAATATCGGTTTCGCAAGTGA	5502		
RESULT 7	AY150274	3540 bp	DNA	circular	BCT 02-JAN-2003
LOCUS	Propionibacterium granulosum	cryptic plasmid pPG01	complete		
DEFINITION	sequence.				
ACCESSION	AY150274				
VERSION	AY150274.1	GI:27465054			
KEYWORDS	Propionibacterium granulosum				
SOURCE	Propionibacterium granulosum				
ORGANISM	Propionibacterium granulosum				

REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Propionibacterineae; Propionibacteriaceae; Propionibacterium.
AUTHORS	1 (bases 1 to 3540)
TITLE	Farrar, M.D. and Holland, K.T.
JOURNAL	Isolation and characterisation of a cryptic plasmid from the human skin commensal Propionibacterium granulosum
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 3540)
TITLE	Farrar, M.D. and Holland, K.T.
JOURNAL	Direct Submission
FEATURES	Submitted (13-SEP-2002) Skin Research Centre, Division of Microbiology, University of Leeds, Leeds LS2 9JT, UK
Location/Qualifiers	Location/Qualifiers
1..3540	
/organism="Propionibacterium granulosum"	
/mol_type="genomic DNA"	
/strain="PF283"	
/db_xref="taxon:33011"	
/plasmid="cryptic plasmid pPG01"	
74..280	
/codon_start=1	
/transl_table=11	
/product="Orf3"	
/protein_id="AA78121.1"	
/db_xref="GI:27465055"	
/translation="MGQYGLNRGFYNGSHNEIIGVPMNGRVIVWPDALLSAMPL	
SSGFSPHVVEPRAEVFHRCADHD"	
449..1792	
/gene="tra"	
449..1792	
/gene="tra"	
/note="plasmid transfer protein"	
/codon_start=1	
/transl_table=11	
/product="Tra"	
/protein_id="AA78122.1"	
/db_xref="GI:27465056"	
/translation="WAGERTVLTGVYWLRLPVLGLLVAFACWCFVSHLVYVIGFA	
VAALAVAFPPSWQVFNQVPMVYVYVWMSLAKAPKPIESYGGDRIVYRPGLOWI	
RSDRHVLVLRVAPGLADSAILEKGAETQQRKSWRTCLIVPAEAGLDLIPR	
DATAGDLPVAPTSWNVPVGVKPDGSEVVDLSHPHSHLVKTRSGKSSFVYGLD	
QMRHLPLVAGVDPDTGILFNELDGWDGADLSKRITNDADAAVQVLSMTDEM	
RRIYLLNCEHRKDRNDFSDPGRLLIIVILEYVPGILERLQNFDSARGARSSDRFA	
SKAALGNRIAEKAGVGVLLVLTQRPDAKIIGPLRAQLTTRVTFQADSDGLRMSH	
PELSBOVKQNSWASGVGFIEADGVIPLTRFSYRAELTDLHRPGASVGVQIDLIQ"	
1979..3280	
/gene="rep"	
1979..3280	
/gene="rep"	
/note="plasmid replication protein"	
/codon_start=1	
/transl_table=11	
/product="Rep"	
/protein_id="AA78123.1"	
/db_xref="GI:27465057"	
/translation="MLNITVTSKAQPHDLCSHEPCCAAARSRORAAAGDRIGTTRAT	
FPVSDSPCKRNEGRHRYEMRDLGNPQVMPLEVRKCGVAPVQSRIALMAGHGAG	
YLAGATCSGWACPVCAKISAHRRDELARVVQVAVGLGFKVMSLTLTQRHAGQDLA	
ELWASLGSWNAVTSRRRWQFECAQLQGVQVWKAVEVTHSGHGHVHVHLVLSKQDP	
TSVDTKIRHRKQRRPTPEYQRPEDFAERWSRGLRKRGVDFIAGSGGLDWOTA	
DSGDEALGRYVAKMNSVDGLANEATLGKFKARNRTPFQILEDLDTGSETDLR	
LWRTVSAHGKALTWKGLRDWAGMESEDFQVAAQDQCGEAVLFDHDAWQIR	
TAGAAFLDLELHGSBGVYAWLKKRRIHYEILPVPWSTST"	

ORIGIN

Query Match	8.3%;	Score 94.4;	DB 1;	Length 3540;
Best Local Similarity	56.2%;	Pred. No. 2.1e-13;		
Matches	198;	Conservative	0;	Mismatches 151; Indels 3; Gaps 1;
QY	181	GGTCCGAAAGTTCTGATTCGGAGCCCTTCGTTCTCGGAAAGGCTGATCTGCCCC	240	
DB	2267	GGTCATGTGTGGTGGCGGTTATGCCGCTTGGCGACGTGCGGAAGTGTGTGGGCTTGCCT	2326	

PUBMED REFERENCE AUTHORS TITLE JOURNAL FEATURES	12948627	2 (bases 1 to 4603)	Tauch,A., Puehler,A., Kalinowski,J. and Thierbach,G. Direct Submission Submitted (05-NOV-2002) Department of Genetics, University of Bielefeld, Universitaetsstrasse 25, Bielefeld D-33615, Germany
	Location/Qualifiers		
	1..4603		
	/organism="Corynebacterium glutamicum"		
	/mol_type="genomic DNA"		
	/strain="22220"		
	/db_xref="taxon:1718"		
	/plasmid="pAG3"		
	complement(200..397)		
	CDS	/note="orf1"	
/codon_start=1			
/evidence=not experimental			
/transl_table=11			
/product="hypothetical protein"			
/protein_id="AAO18193.1"			
/db_xref="GI:27657762"			
/translation="MSKQKLSPPRAKVLAAVDREVTPVQVAHKIGSTQRKAGSALWAL HKANLIDRTFRGIYAPLEVDV"			
complement(613..2061)			
CDS		/gene="repA"	
	complement(613..2061)		
	/note="repA"		
	/note="replication protein; essential for pAG3"		
	/replication		
	/codon_start=1		
	/transl_table=11		
	/product="RepA"		
	/protein_id="AAO18194.1"		
	CDS	/db_xref="GI:27657763"	
/translation="MSILSSKQAEHATSNAPNOATPATHTTGRGLTDTTKHKVKALS PALFRKWLQDRTWKVYTEDTLACGGRVWAGGSVASLVWQAGRARWGLQNSHS WGPSVASVTCRSRANEVDAAIQAWANGAGLTPAHYRGVSTSPDRKORGVSVPV ERGISLMTLLHNSKSLTEVDALAGCQWAVTNAAWRGARTAGSKRYGLAHWY RAIEVTHGNGWHVHLVDFRVLSDVERSLADRVFDRWAARAKVRGMRAPSRDR GIDVVHVAASDDAKSIGYCTCKGMSGTAAETTTGQVTKAEKGNRTTFFQILDLGX QVTKRDHALMEWEGSKGRRTGWSQGTQVILGINSDDQIDDSLGDNDQSEVVAM VGRGAWGATASDEKRAVILDAVATDSAEAGAKAARDVLKLFVGAHTVTAIGLAKTD LTVLPTHDFVKRELHEAVWSGRKQALPVL"			
2396..2785			
/note="orf3"			
/codon_start=1			
/evidence=not experimental			
/transl_table=11			
/product="hypothetical protein"			
/protein_id="AAO18195.1"			
CDS		/db_xref="GI:27657764"	
	/translation="MYRSATCLALHALNVRCFVNVNRETLDMDTYATLNKKNAT KAELEQKDMQOALKQOQALFNAEELEQORNLKPVVTVRVPOYVRDEMKRLAAI QGWTLQOFAANAFAGMIDELSEQDQH"		
	complement(2840..3232)		
	/note="orf4"		
	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18196.1"		
	CDS	/db_xref="GI:27657765"	
/translation="MTKPNVRFNSANDIEAMIGLVIAWNNYCTDKQAKTKAKAIAIA LNKWADELRGRKKDLAHQDETHIERVLHHLFQKATHDNWQLDVEHSRLQSVHR IEAAGEFTRYGNWGEVQLLAETDVE"			
complement(3344..4030)			
/note="orf5"			
/codon_start=1			
/evidence=not experimental			
/transl_table=11			
/product="hypothetical protein"			
/protein_id="AAO18197.1"			
CDS		/db_xref="GI:27657766"	
	/translation="MATPSKPPRGLAEQLEKVRSSLEVIASNVNADHDLSDDGKNNNA WRYTAPHTYIAQVEVALETISIDKAFNDARDKALPTATTSTKLTVVMEQLRII NRGIPESVDGIYKLVTSHEPSPRTALIEELARGRLSDMWVGILSKTSEVAAATE MMQIHARINSVFGNLRMYKALDDRKALFNHWVNTRESADYDMEVPTVTVPPPEKP TNAETVYRAS"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
	/gene="orf6"		
	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
CDS	/protein_id="AAO18198.1"		
	/db_xref="GI:27657767"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
	/gene="orf6"		
	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
CDS	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657768"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
	/gene="orf6"		
	/codon_start=1		
	/evidence=not experimental		
CDS	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657769"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
	/gene="orf6"		
	/codon_start=1		
CDS	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657770"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
	/gene="orf6"		
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657771"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657772"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657773"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657774"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657775"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657776"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657777"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657778"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657779"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657780"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657781"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657782"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657783"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657784"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657785"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657786"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657787"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657788"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657789"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657790"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657791"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657792"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657793"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657794"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657795"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657796"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657797"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657798"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="AAO18199.1"		
	/db_xref="GI:27657799"		
	/translation="MGRNATANWLVQSTQIDNISKALNIDWGDGHVAKANSARAEQL QSAXLELAIDFKRKIASLELSDVLKTDCTTPGERRDQIMIAALSSDKFLKLESTGAPQ EVAGQADAMAALDRFMYATKAAVHGSPDNPNDH"		
	complement(4112..4519)		
	/gene="orf6"		
	complement(4112..4519)		
/gene="orf6"			
CDS	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="hypothetical protein"		

FT -10 signal 410. .415
 FT RBS_ 525. .530
 FT FT 539. .2023
 FT CDS

FEATURES
 source Location/Qualifiers

1. .2297
 /organism="Bifidobacterium breve"
 /mol_type="genomic DNA"
 /db_xref="taxon:1685"

ORIGIN

Query Match 4.5%; Score 51.6; DB 6; Length 2297;
 Best Local Similarity 52.9%; Pred. No. 0.03;
 Matches 111; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 QY 394 GCCTGGAAAGCTGGACCAACGGTGTGGTGGCGTACGGAACGTGAATGTACGGGTGC 453
 Db |||||
 1184 GGCTGGACGAAGATGATTACGGAAGCCCTTGGCAACGGCTCGGAACGTTGGAAATC 1243
 QY 454 GACGGATACGTGGCGGCTGTTGAATCACTACGGAAGAAACGGCTGGCAGTCCAGTT 513
 Db |||||
 1244 AGGGTTTCGTCGGCGGATGGAATCACCTACGTTGTGAACGGCTGGCACCCTCACATT 1303
 QY 514 CACCGCTACTCATGTTTCAGTGGTACGTGAGTGAGAACATCCTCGAATCCTTCTCGGAT 573
 Db |||||
 1304 CATTTCTGTCATGTTCTCGATGGCGATCTGACGATGGCAGCGTGAAGCAATGCAGCA 1363
 QY 574 GCGATGTCGATCGGTGGACTTCCAAATC 603
 Db |||||
 1364 TGGCTGCTGATCGCTGGAACCATGGTC 1393

RESULT 11

AF085719/c
 LOCUS AF085719 5750 bp DNA circular BCT 03-JUN-1999
 DEFINITION Bifidobacterium breve plasmid pCIBb1, complete sequence.
 ACCESSION AF085719
 VERSION AF085719.1 GI:4972586

SOURCE

Bifidobacterium breve
 Bifidobacterium breve
 Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 Bifidobacteriaceae; Bifidobacterium.

1 (bases 1 to 5750)
 O'Riordan, K. and Fitzgerald, G.F.

AUTHORS Molecular characterisation of a 5.75-kb cryptic plasmid from
 Bifidobacterium breve NCFB 2258 and determination of mode of
 replication

JOURNAL FEMS Microbiol. Lett. 174 (2), 285-294 (1999)

MEDLINE 99271179

PUBMED 10339821

REFERENCE 2 (bases 1 to 5750)

O'Riordan, K. and Fitzgerald, G.F.

AUTHORS Direct Submission

TITLE Submitted (22-AUG-1998) Microbiology, University College Cork,

JOURNAL Western Road, Cork, Ireland

FEATURES

source Location/Qualifiers
 1. .5750
 /organism="Bifidobacterium breve"
 /mol_type="genomic DNA"
 /strain="NCFB 2258"
 /db_xref="taxon:1685"
 /plasmid="pCIBb1"
 1463. .2062
 /gene="par"
 /note="partitioning gene"
 1463. .2062
 /gene="par"
 /note="par protein; putative involvement in plasmid
 partition"
 /codon_start=1
 /transl_table=11
 /product="partitioning protein"

gene

CDS

/protein_id="AAD34708.1"
 /db_xref="GI:4972587"
 translation="MILSIITASKGKTGKTSSIFLACAAHTADPSLSITLLDAPQOS
 ATWAEEAQNKLPLPFWKVAGSEIENLKNELGLTIDTAPGFTDLTQIIAKSDL
 ILITTSPTQIDARAWKTFFAIOIRAGGTSAILITRANTRTKDFRDMIDDLNDNAN
 ILFTVIRQSTRYSKPFGHTPKTYDYADAWQEIQKAGE"
 complement(2373. .3515)
 /gene="rep"
 /note="replication gene"
 complement(2373. .3515)
 /gene="rep"
 /note="rep"
 /note="rep; similar to replication proteins of
 streptomyces plasmids and other members of the pC194
 family that replicate by a rolling circle mechanism"
 /codon_start=1
 /transl_table=11
 /product="replication protein"
 /protein_id="AAD34709.1"
 /db_xref="GI:4972588"

/translation="MTNGEKARFTGMLCGSIWACPTCSAIIHERAHEVALAIGNHA
 EKLKAAADQWAEHQRLPELMSVDSFGNYIFGLTLRHDRTPMLATLDAIKG
 WTKMNGSPQWASERKIRGFVRAIETTVNGVNGHPHIFVMEFDGLDDGOREAMQ
 QWLLDRWTKVQKAKYKKDGNPNVAVNDEHGIDLOFKSGKDACTAAAEYITKIQ
 GDKGVTLAEIARGDINKRGMSPVNPOLLDSGLGLSDFQREDLMLEWQATLRRR
 CITWSRGLKDEMEVEELEDELAEKADELFGLVYVNPVNRVYKDIRKASPETLLADLD
 AAREDDQEVARLLPFGSVILTDQQAIDAEAKPGDYLTMTSMVMV"
 complement(3926. .3934)
 /note="region conforming to double stranded origin;
 nicking site of rolling circle plasmids"
 complement(3985. .4845)
 /gene="ftsK gene homolog"
 complement(3985. .4845)
 /gene="ftsK gene homolog"
 /note="hypothetical protein; similar to transfer proteins
 and FtsK homologs"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAD34710.1"
 /db_xref="GI:4972589"

rep_origin

/translation="MKWPNHNKPEPLPRLPFPDPATGTTVGLANGNATLIYV
 NTSGVVVGIPGSGKTAGWVIVLALYLSGCCNHHVLDGGDDWGFSHATTFVVG
 DLDVHDTLLRLDDMKCRISMQRYSANYVWQDKRPPLEVIIIDSCQSFNNAK
 GILGKPAKABEITAAATEIVOKGRSGGFLFAITQKPTDLSFSALENCENRVC
 FRYKTPAARAVILGMDPDSGSPSTPIPPARRGAIIGLATGEDVMCRFAYVSEAEER
 AVVANHKGVI"

ORIGIN

Query Match 4.5%; Score 51.6; DB 1; Length 5750;
 Best Local Similarity 52.9%; Pred. No. 0.029;
 Matches 111; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 QY 394 GCCTGGAAAGCTGGACCAACGGTGTGGTGGCGTACGGAACGTGAATGTACGGTGC 453
 Db |||||
 3212 GGCTGGACGAAGATGATTACGGAAGCCCTTGGCAACGGCCCTCGGAACGTTGGAATC 3153
 QY 454 GACGATACGTGGCGCGTGTGAAATCACTACGGAAGAAACGGCTGGCAGTCCAGTT 513
 Db |||||
 3152 AGGGGTTTCGTCGGCGGATTTGAAATCACTACGTTGTAACGGCTGGCACCCTCACATT 3093
 QY 514 CACGCGCTACTCATGTTTCAGTGGTACGTGAGTGAACATCTCTCGAATCCTTCTCGGAT 573
 Db |||||
 3092 CATTTCTGTCATGTTTCTCGATGGCGATCTGGACGATGGGCGAGCGTGAGCAATGCAGCA 3033
 QY 574 GCGATGTTCCATCGGTGGACTTCCAAATC 603
 Db |||||
 3032 TGGCTGCTCATCGCTGGAAACCATGGTC 3003

RESULT 12

AF429315/c
 LOCUS AF429315 125020 bp DNA linear PRI 18-JAN-2002
 DEFINITION Homo sapiens junctionalillin 3 (JPH3) gene, partial cds.


```
/mol_type="genomic DNA"
/db_xref="taxon:1744"
/lab_host="DPF2"
/plasmid="pLMB108"
complement (223. .1434)
/gene="rep"
complement (223. .1434)
/gene="rep"
/codon_start=1
/transl_table=11
/product="putative rep protein"
/protein_id="CAA07175.1"
/db_xref="GI:3212128"
/db_xref="COA:069764"
/db_xref="TREMBL:O69764"
/translation="MVVDARARSGVRCGGGSPAVPLGALDNHANNVSPSAVPA
GEVADKRRFESVRYLWRLTSLKRVAFGRVAASAVASVGRCDRAGFAGLQSC
SVWACPVCAKIMARGLGELGAAVEVTWKHGRVAFMTFVRSRCDLSLTAVWDGVA
SGWRVTSKGKWTSDQRLHGVGVFVRVETVTHGRGWHVHLVFLVFLVGFGLALAH
RSMFRWEGVLAAGLCTPLAARDVOQMSNAATGLDLHLYLSKAOFGKIGHELINS
QKTSARSLSTRTWELGDAAKLEKGVLEWFEWEGSRGRROIGWSAGDRRLGLM
VESDDVIAAEVGSVADTVALLITGDGWRRLVGQOKLYECLRSCELGQGLARWQGE
HGIEHELVEVD"
```

ORIGIN

```
Query Match 4.1%; Score 47.2; DB 1; Length 2051;
Best Local Similarity 49.4%; Pred. No. 0.42; 178; Indels 3; Gaps 2;
Matches 177; Conservative 0; Mismatches 178; Indels 3; Gaps 2;

QY 194 CTGATTTCGGAGGCTTCGTTCTGCGGAAAGGCTGATCGCCCTGCTGTCGGCGAA 253
DB 1154 CCGGTTTCGTTGCTGCAGAGCTGCGATCGTCTGCGGCTGCGCGGTGTAACCGA 1095

QY 254 AAGTCGTGCATCTGTCGACAGCAATTTCTCAAGTTGTG-CTCATCACTCGGAC- 311
DB 1094 AGATCATGGCTCGCGTGGCCCTTGAACCTCGCGCGCGGCTCGAGACCTGGACTAAGCACG 1035

QY 312 -TGATCTGTTGCGATGTCGACGATGACCATGCGCCATACAGCTGGTCAGCGGCTCCAG 370
DB 1034 GTGTCGTGTGGCTTCATGACGTTTCAGGTTGCGGCTGCGGCTGCTGTAAGGATTCCTTGACGG 975

QY 371 ACCTATGACTGGACTTTCGGCAGCCTCGGAAGCTGCGACCAACGCTGCTGTGGCGTA 430
DB 974 CTGTTGGGATGTGTGCTTCGGGGTGGCGGAGTCACTTCGGGCAAGGCTGGAGCT 915

QY 431 CGGAAGCTGAATGATCGGCTGCGACGAGTACGTCGGCGCTGTTGAATCACTCACGAA 490
DB 914 CTGACCAAGCTCCGTCACGGGGTCGAGGCTTCGTGCGTGTGCTGAGTTACGCACGGC 855

QY 491 AAAACGGCTGGCAGCTCCAGCTTTCACGCGCTACTCATGTTTCAGTGGTGAGTGATGA 548
DB 854 GCAATGGGTGGCATGTGCACTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGA 797
```

RESULT 14

```
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 125020)
REFERENCE Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
AUTHORS Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE 21583737
```

PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

```
11694876
2 (bases 1 to 125020)
Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Direct Submission
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
Location/Qualifiers
1. .125020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16Q24.3; between D16S520 and WI-12410"
/notes="isolated from a patient with Huntington's
Disease-like 2 (HDL2)"
complement (35581. .35746)
Disease-like 2 (HDL2)"
/rpt_type=tandem
/rpt_unit="ctg"
complement (<36507. .>36887)
/gene="JPH3"
/notes="synonym: JP3"
complement (<36507. .>36887)
/gene="JPH3"
/product="junctophilin 3"
complement (<36507. .36887)
/gene="JPH3"
/notes="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AA140941.1"
/db_xref="GI:17646245"
/translators="MSSGRRFDDGGSCYCGWEDKAGHGVCTGPKGGGYTGWS
HGFEVLGVYTPSGNTYQGTAAQKRGHIGLESKGWVYKGEWTHGFKRGYVRECA
NGAKYEGTWSNGLQDGYGTYSDG"
```

ORIGIN

```
Query Match 4.1%; Score 46.4; DB 9; Length 125020;
Best Local Similarity 10.8%; Pred. No. 0.61;
Matches 95; Conservative 397; Mismatches 378; Indels 6; Gaps 3;

QY 108 AATCACACCTCAGAACATTTAAAGCCTGTGCGCGCGCATTTCTGCGGTGAACGGTGT 167
DB 50582 RWTGKBGWSKSHSCHRRTWTDYRMKYTYCCCYTAYMBBYHYCYDBGWSGWSGGW 50641

QY 168 GACCATTTGCAACGGTCCGAAAGTTCTGATTCGAGGCGCTTCGTTCTCTCGGAAAGGG 227
DB 50642 RGMKMBDWBMBSSMMKSMGMDGTDKDYWCYSSMSVDVVBWRWBDTSWGSBSHYMEGVH 50701

QY 228 CTGATCTGCCCTGCTGTGCGGGAAGAAAGTCGGTGCACATCGTGCAGACGAAATTTCTCA 287
DB 50702 TGYMRWVGYGBCDTHVTYDGGHSGWMMKSCGYSDBGKCAKBSKMKCTGSMTKYTCGY 50761

QY 288 AGTTGTTGTCATCAACTCGGAGCTGATCTGTTGGATGGTGCACGATGACCATGCGCCA 347
DB 50762 GKTWSSKSDBSYYSBSHYBWSYMCCKRKBTSASTDYRYKBTGYRIVRSCSWYVK 50821

QY 348 TACAGCTGTGTCAGCGCTCCA-CGACCTATGAGCTGACCTTTTCGGCAGCTGGAAGAGCTG 406
DB 50822 YBSMSVSARKSGHKVRBSSMKCSRRKDMSTSSMWSRSMKDRKCSRSHSHYSKMGKWT 50881

QY 407 CGACCAACGGTCTCGTTCGGCTACGGAAGCTGAATGTACGGCTGCGACGATGATGTCG 466
DB 50882 DVDRCYACSBMYKVKHKKHYKDDSHBSWGSHMKMGAWRVSRMSRCSVHDSMVD 50941

QY 467 GCGCTGTGAAATCACTACGGAAGAAACGGCTGGCAGCTCCACGTTCCAGCGCTACTCA 526
DB 50942 KSRMRMGAMRGMKCYCYTSMRSTKSRMSKMSWRWGSKC----YCYGWSCTKMRSM 50997

QY 527 TGTTCAGTGTGATCGTGAAGTGAACATCTCGAATCTTCTCGGATCGGATGTCGATC 586
DB 50998 GCSYSTGSSMKCMGWYCCWAGRSRKYCKSGSRMSWAGSMASMSKMGKWTSGMSGRCC 51057
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2004, 22:12:19 ; Search time 637 Seconds

(without alignments)
9394.568 Million cell updates/sec

Title: US-10-007-527A-1

Perfect score: 1140

Sequence: 1 atgaccagcgtaagtgtga.....taatatcggttcgaagtga 1140

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	100.0	1140	6	ABQ76122
C 2	1140	100.0	6334	6	ABQ76124 Rhodococc
C 3	1140	100.0	9652	6	ABQ76126 Plasmid p
C 4	1140	100.0	11241	6	ABQ76125 Plasmid p
C 5	1140	100.0	11241	12	ADH10182 E. coli-R
6	51.6	4.5	2297	2	AAV58945
7	44.4	3.9	2000	8	ADA71938 Rice gene
8	40.6	3.6	1189	10	ACH00861
9	40.6	3.6	1244	10	ACH00860
10	40.6	3.6	1841	8	ACC00046
11	37.4	3.3	2708	8	ACC00048
12	37.4	3.3	4447	3	AAA90951
13	37.4	3.3	4447	3	AAA90934
14	37.4	3.3	4447	6	ABL49734
15	37.4	3.3	4447	6	ABL49733
16	37.4	3.3	4447	6	ADL22583
17	37.4	3.3	4447	6	ADL22582
18	37.4	3.3	4447	10	ADB66212
C 19	36.6	3.2	2000	8	ADA71779
C 20	36.4	3.2	2000	8	ADA71938
C 21	34.8	3.1	498	8	ACA18944

C 22	34.8	3.1	702	8	ACA14012
23	34.8	3.1	861	8	ACA53382
24	34.8	3.1	1631	3	AAA66016
C 25	34.8	3.1	4590	5	AHH24065
26	34.6	3.0	1505	2	AAQ55750
C 27	34.6	3.0	1954	6	ABI99802
C 28	34.4	3.0	1173	5	AAF26431
C 29	34.2	3.0	684	4	AAF26442
30	34.2	3.0	1314	6	ABQ90222
C 31	34.2	3.0	3399	3	ADC78561
32	34.2	3.0	3448	4	ADP26441
33	34.2	3.0	4438	4	AS559525
C 34	34.2	3.0	4438	8	ACF64454
C 35	34	3.0	534	6	ABQ41452
36	34	3.0	534	6	ABQ41452
37	34	3.0	1137	8	ACD05604
C 38	33.8	2.9	2019	12	ADQ22781
39	33.6	2.9	786	8	ACA00393
40	33.6	2.9	832	6	ABQ20710
C 41	33.6	2.9	832	6	ABQ20711
C 42	33.6	2.9	1134	3	AAA58791
C 43	33.6	2.9	1134	6	AAQ39776
C 44	33.6	2.9	1134	6	ABS53140
C 45	33.6	2.9	1281	8	ACA39926

ALIGNMENTS

RESULT 1
ABQ76122
ID ABQ76122 standard; DNA; 1140 BP.
XX
AC ABQ76122;
XX
DT 13-JAN-2003 (first entry)
XX
DE Rhodococcus AN12 replication protein Rep DNA.
XX
KW Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase; gene;
KW shuttle vector; Rep; db.
XX
OS Rhodococcus erythropolis.
XX
PN WO200255709-A2.
XX
PD 18-JUL-2002.
XX
PF 12-DEC-2001; 2001WO-US047868.
XX
PR 12-DEC-2000; 2000US-0254868P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX
DR WPI; 2002-557827/59.
XX
DR P-PSDB; ABB84278.
XX
PT New nucleic acid molecule encoding replication protein/plasmid stability
XX protein, useful in cloning and expression vectors, particularly shuttle
XX vectors for expression of heterologous genes in Rhodococcus species.
XX
PS Claim 2; Page 63-64; 96pp; English.
XX
CC This invention describes a novel nucleic acid encoding a replication
XX protein or a plasmid stability protein. The product of the invention is
XX useful for expression of nucleic acid such as genes encoding enzymes
XX involved in the production of isoprenoid molecules, polyhydroxyalkanoic
XX acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile

CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence encodes the Rhodococcus AN12 Rep protein
CC described in the disclosure of the invention
XX
SQ Sequence 1140 BP; 241 A; 287 C; 347 G; 265 T; 0 U; 0 Other;
Query Match 100.0%; Score 1140; DB 6; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACCAAGCTGTAAGTCTGAACACCTTTCCGGCAAGACCGCGCTCCGCTGCTGTCG 60
DB 1 ATGACCAAGCTGTAAGTCTGAACACCTTTCCGGCAAGACCGCGCTCCGCTGCTGTCG 60
QY 61 TCCGATAGCGCGGATCCGGACGAACTGGACCCAACTTCAACAAATCACCACGTCA 120
DB 61 TCCGATAGCGCGGATCCGGACGAACTGGACCCAACTTCAACAAATCACCACGTCA 120
QY 121 GAAACATTTAAGCCTGTGGCGCGGATTTCTGGCGTGAACGGTGTGACCAATTGTCAAC 180
DB 121 GAAACATTTAAGCCTGTGGCGCGGATTTCTGGCGTGAACGGTGTGACCAATTGTCAAC 180
QY 181 GGTCCGAAAGTTCTGGATTGGAGGCTTGGAGGCTTCTTCTCGGAAAGGCTGGATCTGCC 240
DB 181 GGTCCGAAAGTTCTGGATTGGAGGCTTGGAGGCTTCTTCTCGGAAAGGCTGGATCTGCC 240
QY 241 TGCTGTGGGAAAGTCTGGTGCATCTGTCAGACGAAATTTCTCAAGTTGTTCCTCAT 300
DB 241 TGCTGTGGGAAAGTCTGGTGCATCTGTCAGACGAAATTTCTCAAGTTGTTCCTCAT 300
QY 301 CAACCTCGGACCTGGATCTGTTGCGATGTGACGATGACCATCGCCATACAGCTGGTCAG 360
DB 301 CAACCTCGGACCTGGATCTGTTGCGATGTGACGATGACCATCGCCATACAGCTGGTCAG 360
QY 361 CGGCTCCACGACCTATGACTGGAATCTTTCGGCAGCTTGGAAAGCTCGCAACACGCTGT 420
DB 361 CGGCTCCACGACCTATGACTGGAATCTTTCGGCAGCTTGGAAAGCTCGCAACACGCTGT 420
QY 421 CGTTGGCGTACGGAACGTAAGTGTACGCTCGGACGATGATCGCGCTGTGAAATC 480
DB 421 CGTTGGCGTACGGAACGTAAGTGTACGCTCGGACGATGATCGCGCTGTGAAATC 480
QY 481 ACTCACGGAAGAAACGGCTGGCAGCTCCACGTTCCAGCTACTCATGTTCAGTGGTCAG 540
DB 481 ACTCACGGAAGAAACGGCTGGCAGCTCCACGTTCCAGCTACTCATGTTCAGTGGTCAG 540
QY 541 GTGAGTGAACATCTCGAATCTTCTCGGATGCGATGTTGCGATCGTGGGACTTCCAAA 600
DB 541 GTGAGTGAACATCTCGAATCTTCTCGGATGCGATGTTGCGATCGTGGGACTTCCAAA 600
QY 601 CTCGTATCTCTCGGATCTGCTGGCCACTACGTAATTCGGGTGCTCGATGTCGAAAG 660
DB 601 CTCGTATCTCTCGGATCTGCTGGCCACTACGTAATTCGGGTGCTCGATGTCGAAAG 660
QY 661 ATCCGCGGTGAAGCTGATCAAGTTCCTCGGTGATCTGACGAAATTCATCTGCGCT 720
DB 661 ATCCGCGGTGAAGCTGATCAAGTTCCTCGGTGATCTGACGAAATTCATCTGCGCT 720
QY 721 GGTATGGAGGTTGGTATGGCGGAAAGAGTGTGACATGGCAACCGTGCACCTGG 780
DB 721 GGTATGGAGGTTGGTATGGCGGAAAGAGTGTGACATGGCAACCGTGCACCTGG 780
QY 781 GAAATCGCTGTTGATGTCAGTGGCGGGATCCCAAGCGTTGGAATCTGCGGAGATTT 840
DB 781 GAAATCGCTGTTGATGTCAGTGGCGGGATCCCAAGCGTTGGAATCTGCGGAGATTT 840
QY 841 GAGTTTGGTTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTGCGTCCCGAGCT 900
DB 841 GAGTTTGGTTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTGCGTCCCGAGCT 900

QY 901 GGTCTTGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC 960
DB 901 GGTCTTGGCGAGAACTAAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC 960
QY 961 ATGCTTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTGTGCGCTTACGTC 1020
DB 961 ATGCTTTCGATCATTTCCGGCGCGATCGTGGATGATGATTCGGACTTGTGCGCTTACGTC 1020
QY 1021 TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAATCTTCGTGATCAC 1080
DB 1021 TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAATCTTCGTGATCAC 1080
QY 1081 TTGCATTATCGATTGCCCGCGAGCGATGTGCGGCCCGGATAATATCGTTTCGCAAGTGA 1140
DB 1081 TTGCATTATCGATTGCCCGCGAGCGATGTGCGGCCCGGATAATATCGTTTCGCAAGTGA 1140

RESULT 2

ABQ76124/C
ID ABQ76124 standard; DNA; 6334 BP.
XX
AC ABQ76124;
XX
DT 13-JAN-2003 (first entry)
XX
DE Rhodococcus AN12 derived plasmid pAN12 DNA.
XX
KW Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW shuttle vector; circular; ds.
XX
OS Rhodococcus erythropolis.
OS Synthetic.
XX
PN WO200255709-A2.
XX
PD 18-JUL-2002.
XX
PF 12-DEC-2001; 2001WO-US047868.
XX
PR 12-DEC-2000; 2000US-0254868P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX
DR WPI; 2002-557827/59.
XX
PT New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.
XX
PS Claim 19; Page 68-71; 95pp; English.
XX
CC This invention describes a novel nucleic acid encoding a replication
CC protein or a plasmid stability protein. The product of the invention is
CC useful for expression of nucleic acid such as genes encoding enzymes
CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence represents the Rhodococcus AN12 derived
CC plasmid pAN12 DNA described in the disclosure of the invention
XX
SQ Sequence 6334 BP; 1386 A; 1807 C; 1681 G; 1460 T; 0 U; 0 Other;

Query Match 100.0%; Score 1140; DB 6; Length 6334;

Best Local Similarity 100.0%; Pred. No. 0;		Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGACACGCTAAGTCTGAACACCTTTCCGCGAAAGACCGGCTCCCGTCTCGTGTGCG	60
Db	3051	ATGACACGCTAAGTCTGAACACCTTTCCGCGAAAGACCGGCTCCCGTCTCGTGTGCG	2992
QY	61	TCCGATAGCCGGCATCCGCGACGAACTCGGACCCAAACTTCAACAATCACCACGTCA	120
Db	2991	TCCGATAGCCGGCATCCGCGACGAACTCGGACCCAAACTTCAACAATCACCACGTCA	2932
QY	121	GAACATTTAAGCCCTGTGGCGCGCGATTTCTGCGGTGAACGCGTGTGACCAATGTCAC	180
Db	2931	GAACATTTAAGCCCTGTGGCGCGCGATTTCTGCGGTGAACGCGTGTGACCAATGTCAC	2872
QY	181	GGTCGGAAGGTTCTGGATTGGAGGCTTCCTCGGAAAGGGCTGGATCTGCCCC	240
Db	2871	GGTCGGAAGGTTCTGGATTGGAGGCTTCCTCGGAAAGGGCTGGATCTGCCCC	2812
QY	241	TGCTGTGCGGAAAGTCGTGCACATCGTCGACGAAATTTCTCAAGTTGTTCTCAT	300
Db	2811	TGCTGTGCGGAAAGTCGTGCACATCGTCGACGAAATTTCTCAAGTTGTTCTCAT	2752
QY	301	CAACTCGGCACTGGATCTGTTGCGATGTGACATGCCGCATACAGCTGCTCAG	360
Db	2751	CAACTCGGCACTGGATCTGTTGCGATGTGACATGCCGCATACAGCTGCTCAG	2692
QY	361	CGGCTCCACGACTTANGACTGGACTTTTCGGCAGCTCGGAAAGCTGCGACCAACGGT	420
Db	2691	CGGCTCCACGACTTANGACTGGACTTTTCGGCAGCTCGGAAAGCTGCGACCAACGGT	2632
QY	421	CGTTGGCGTACGGAACGTGAATGTACGGCTCGGACGATAGCTGCGGCTGTTGAAATC	480
Db	2631	CGTTGGCGTACGGAACGTGAATGTACGGCTCGGACGATAGCTGCGGCTGTTGAAATC	2572
QY	481	ACTCAGGAAAAACGGCTGGCAGCTCCACGCTTCCAGCGTACTCATGTTTCAAGTGTGAC	540
Db	2571	ACTCAGGAAAAACGGCTGGCAGCTCCACGCTTCCAGCGTACTCATGTTTCAAGTGTGAC	2512
QY	541	GTGAGTGAAACATCCTCGAATCTTCTCGGATGCGATGTTTCGATCGGTGGATTTCCAAA	600
Db	2511	GTGAGTGAAACATCCTCGAATCTTCTCGGATGCGATGTTTCGATCGGTGGATTTCCAAA	2452
QY	601	CTCGTATCTCGGATTTGCTCGGCACTACGTAATTCGGTGGTCTCGATGTACGAAG	660
Db	2451	CTCGTATCTCGGATTTGCTCGGCACTACGTAATTCGGTGGTCTCGATGTACGAAG	2392
QY	661	ATCGGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAAAATTCATCTGGCGTT	720
Db	2391	ATCGGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAAAATTCATCTGGCGTT	2332
QY	721	GGTATGAGGTTGTGAGTGGCGACGGAAGAAAGTGTGACATGGCAACCGTGACCCCTGG	780
Db	2331	GGTATGAGGTTGTGAGTGGCGACGGAAGAAAGTGTGACATGGCAACCGTGACCCCTGG	2272
QY	781	GAATCGGTGTGATGCTAGTGGCGGATCCACAGCGTTGGAATGTGGGAGAAATTT	840
Db	2271	GAATCGGTGTGATGCTAGTGGCGGATCCACAGCGTTGGAATGTGGGAGAAATTT	2212
QY	841	GAGTTTGGTTCGATGGGAGCTCGGCAATTCGGTGGTTCGGTGGATTCGGTCCCGAGCT	900
Db	2211	GAGTTTGGTTCGATGGGAGCTCGGCAATTCGGTGGTTCGGTGGATTCGGTCCCGAGCT	2152
QY	901	GGCTTGGGCGAAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAATTCGCCCGGTC	960
Db	2151	GGCTTGGGCGAAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAATTCGCCCGGTC	2092
QY	961	ATGGTTCGATCATTTCCGCGCGATCGTGGATGATTCGGAATTTGCGGCTTACGTC	1020
Db	2091	ATGGTTCGATCATTTCCGCGCGATCGTGGATGATTCGGAATTTGCGGCTTACGTC	2032
QY	1021	TTCCGCGGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGTGATCAC	1080

Db	2031	TTCCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGAAAAATCTTCGTGATCAC	1972
QY	1081	TTGCATTATCGATTGCCCGCAGCGGATGTGGGCCCGGATTAATATCGGTTCCGAAAGTCA	1140
Db	1971	TTGCATTATCGATTGCCCGCAGCGGATGTGGGCCCGGATTAATATCGGTTCCGAAAGTCA	1912
RESULT 3			
ABQ76126/c			
ID	ABQ76126	standard; DNA; 9652 BP.	
XX	AC	ABQ76126;	
XX	AC	ABQ76126;	
DT	13-JAN-2003	(first entry)	
XX	Plasmid	prHBR171 DNA.	
DE	DE		
XX	Plasmid	stability protein; replication protein; ethylene forming enzyme;	
KW	carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;		
KW	polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;		
KW	alcohol dehydrogenase; terpene synthase; cholesterol oxidase;		
KW	shuttle vector; circular; ds.		
OS	Synthetic.		
XX	WO200255709-A2.		
PN	18-JUL-2002.		
XX	12-DEC-2001; 2001WO-US047868.		
PF	12-DEC-2001; 2000US-0254868P.		
PR	(DUPO) DU PONT DE NEMOURS & CO E I.		
XX	Bramucci MG, Cheng Q, Kostichka KN, Tomb J;		
PI	WPI; 2002-557827/59.		
DR	New nucleic acid molecule encoding replication protein/plasmid stability		
XX	protein, useful in cloning and expression vectors, particularly shuttle		
PT	vectors for expression of heterologous genes in Rhodococcus species.		
PS	Claim 27; Page 72; 96pp; English.		
CC	This invention describes a novel nucleic acid encoding a replication		
CC	protein or a plasmid stability protein. The product of the invention is		
CC	useful for expression of nucleic acid such as genes encoding enzymes		
CC	involved in the production of isoprenoid molecules, polyhydroxyalkanoic		
CC	acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile		
CC	hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol		
CC	dehydrogenase, terpene synthases, and cholesterol oxidase in an		
CC	Actinomycetales bacteria. The replication protein or plasmid stability		
CC	protein are useful in cloning and expression vectors and particularly in		
CC	shuttle vectors for the expression of homologous and heterologous genes		
CC	in Rhodococcus sp. This sequence represents the Plasmid prHBR171 DNA		
CC	described in the disclosure of the invention		
XX	Sequence 9652 BP; 2097 A; 2729 C; 2559 G; 2267 T; 0 U; 0 Other;		
Query Match 100.0%; Score 1140; DB 6; Length 9652;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGACACGCTAAGTCTGAACACCTTTCCGCGAAAGACCGGCTCCCGTCTCGTGTGCG	60
Db	5052	ATGACACGCTAAGTCTGAACACCTTTCCGCGAAAGACCGGCTCCCGTCTCGTGTGCG	4993
QY	61	TCCGATAGCCGGCATCCGCGACGAACTCGGACCCAAACTTCAACAATCACCACGTCA	120
Db	4992	TCCGATAGCCGGCATCCGCGACGAACTCGGACCCAAACTTCAACAATCACCACGTCA	4933
QY	121	GAACATTTAAGCTGTGGCGCGGATTTCTGGCGTGAACGGTGTGACCATGTCAC	180

```
Db 4932 GAAACATTTAAACGCTTGGCCGCGGATTTCTGGGTGAACGGTGTGACCAATGTCAAC 4873
QY 181 GGTCCGAAAGTTCTGGATTCGGAGCCCTTCCTTCGCGAAAGGGCTGATCTGCCCC 240
Db 4872 GGTCCGAAAGTTCTGGATTCGGAGCCCTTCCTTCGCGAAAGGGCTGATCTGCCCC 4813
QY 241 TCGTGTGGGGAAGTGGTGCATCTGTCAGAGAAATTTCTCAAGTTGTTGCTCAT 300
Db 4812 TCGTGTGGGGAAGTGGTGCATCTGTCAGAGAAATTTCTCAAGTTGTTGCTCAT 4753
QY 301 CAACTCGGAGCTGGATCTGTCGATGTGTGACATGACCATCGCCATACAGCTGTCTAG 360
Db 4752 CAACTCGGAGCTGGATCTGTCGATGTGTGACATGACCATCGCCATACAGCTGTCTAG 4693
QY 361 CGGCTCCAGCACTATGGAATCTGCGAGCTTCGCGAGCTGGAAGCTGCGACCAACGCTGT 420
Db 4692 CGGCTCCAGCACTATGGAATCTGCGAGCTTCGCGAGCTGGAAGCTGCGACCAACGCTGT 4633
QY 421 CGTTCGGCTACGGAAGCTGGAATGTACGCTGCGAGCGATACGTGGCGCTGTTGAAATC 480
Db 4632 CGTTCGGCTACGGAAGCTGGAATGTACGCTGCGAGCGATACGTGGCGCTGTTGAAATC 4573
QY 481 ACTCACGGAAGAAACGGCTGGCAGCTGCCACGTTCCAGCGCTACTCATGTTCAAGTGTGTAG 540
Db 4572 ACTCACGGAAGAAACGGCTGGCAGCTGCCACGTTCCAGCGCTACTCATGTTCAAGTGTGTAG 4513
QY 541 GTGAGTGAAGCAATCTCGAATCTTCTCGATGCGATGTTGATCGGTGGAATCTCCAAA 600
Db 4512 GTGAGTGAAGCAATCTCGAATCTTCTCGATGCGATGTTGATCGGTGGAATCTCCAAA 4453
QY 601 CTCGTATCTCTGGATTTCTCGCCGCACTACGTAATTCGGTGGTCTCGATCTAGCAAG 660
Db 4452 CTCGTATCTCTGGATTTCTCGCCGCACTACGTAATTCGGTGGTCTCGATCTAGCAAG 4393
QY 661 ATCGCGGTGGAAGCTGATCAAGTTCTCGCTGGTATCTGACGAAATTCATCTGCGCT 720
Db 4392 ATCGCGGTGGAAGCTGATCAAGTTCTCGCTGGTATCTGACGAAATTCATCTGCGCT 4333
QY 721 GGTATGGAGTTGGTAGTGGGACGAGAAAGTGGTGCATGCGCAACCGTGACCCCTGG 780
Db 4332 GGTATGGAGTTGGTAGTGGGACGAGAAAGTGGTGCATGCGCAACCGTGACCCCTGG 4273
QY 781 GAAATCGCTGTTGATGCACTGGCGGGATCCACAGCGTTGGAATCTGGGCGAGATTT 840
Db 4272 GAAATCGCTGTTGATGCACTGGCGGGATCCACAGCGTTGGAATCTGGGCGAGATTT 4213
QY 841 GAGTTTGGTTGATGGGAGCTCGGCAATCGCGTGGTCCCGTGGATTGCGTCCCGAGCT 900
Db 4212 GAGTTTGGTTGATGGGAGCTCGGCAATCGCGTGGTCCCGTGGATTGCGTCCCGAGCT 4153
QY 901 GGTCTGGGGGAGAACTAAGATGCTAGATGCTGTTGAGCAGAGAAATCTGCCCGCGTC 960
Db 4152 GGTCTGGGGGAGAACTAAGATGCTAGATGCTGTTGAGCAGAGAAATCTGCCCGCGTC 4093
QY 961 ATGTTTGGCATCATTCGCGCGGATCGTGTGATGATGATTCGGACTTGTGCGCTTACGTC 1020
Db 4092 ATGTTTGGCATCATTCGCGCGGATCGTGTGATGATGATTCGGACTTGTGCGCTTACGTC 4033
QY 1021 TTCGCGAGATCCTCGGACTCGTGAAGCTGCGCGGACTTGGGAAATCTTCGTGATCAC 1080
Db 4032 TTCGCGAGATCCTCGGACTCGTGAAGCTGCGCGGACTTGGGAAATCTTCGTGATCAC 3973
QY 1081 TTGCATTATCGATTGCCCGAGGGATGTGCGGCCCCCGCATATATCGGTTGCGAAGTGA 1140
Db 3972 TTGCATTATCGATTGCCCGAGGGATGTGCGGCCCCCGCATATATCGGTTGCGAAGTGA 3913
```

RESULT 4

ABQ76125/c

ID ABQ76125 standard; DNA; 11241 BP.

XX

AC ABQ76125;

```
Db 6341 CAACTCGGACTGATCTGTTGGATGTTGACGATGACATGCGCCATACAGCTGTCAG 6282
Qy 361 CGGCTCCACGACCTATGAGTGGACTTTTCGGCAGCCTCGAAAGCTGCGACCAACGGTCTG 420
Db 6281 CGGCTCCACGACCTATGAGTGGACTTTTCGGCAGCCTCGAAAGCTGCGACCAACGGTCTG 6222
Qy 421 CGTTGGCGTACGGACGTTGAATGTACGGCTGCGAGCGGATACGTTGGCGCTGTTGAATC 480
Db 6221 CGTTGGCGTACGGACGTTGAATGTACGGCTGCGAGCGGATACGTTGGCGCTGTTGAATC 6162
Qy 481 ACTCAGGAAAAAAGCGGTGGCAGCTCCACGCTTACGCGCTACTCATGTTCAAGTGGTAC 540
Db 6161 ACTCAGGAAAAAAGCGGTGGCAGCTCCACGCTTACGCGCTACTCATGTTCAAGTGGTAC 6102
Qy 541 GTGAGTGAACATCCTCGAATCCTTCGATGCGATGTTTCGATGCGGTGACATTCGAAA 600
Db 6101 GTGAGTGAACATCCTCGAATCCTTCGATGCGATGTTTCGATGCGGTGACATTCGAAA 6042
Qy 601 CTCGTATCTCTGGGATTTGCTGGCCACTACGTAATTCGGTGGTCTCGATGTAGGAAAG 660
Db 6041 CTCGTATCTCTGGGATTTGCTGGCCACTACGTAATTCGGTGGTCTCGATGTAGGAAAG 5982
Qy 661 ATCGGCGGTGAAGCTGATCAAGTTCCTCGCTGCTATCTGACGAAAAATTCGATCTGGCTT 720
Db 5981 ATCGGCGGTGAAGCTGATCAAGTTCCTCGCTGCTATCTGACGAAAAATTCGATCTGGCTT 5922
Qy 721 GGTATGAGGTTGTGATGGCGACGGAAGAAAGTGTGACATGCGAACCGTGCACCTGG 780
Db 5921 GGTATGAGGTTGTGATGGCGACGGAAGAAAGTGTGACATGCGAACCGTGCACCTGG 5862
Qy 781 GAAATCGCTGTTGATGAGTGGCGGGGATCCACAGGTTTGGAACTGTGCGGAGATTT 840
Db 5861 GAAATCGCTGTTGATGAGTGGCGGGGATCCACAGGTTTGGAACTGTGCGGAGATTT 5802
Qy 841 GAGTTTGGTTCGATGGGACGTGCGGCAATCGCTGCTCGGATTCGCTGCGGAGCT 900
Db 5801 GAGTTTGGTTCGATGGGACGTGCGGCAATCGCTGCTCGGATTCGCTGCGGAGCT 5742
Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 5682
Qy 961 ATGGTTCGATCANTCGGCGCGATCGTGGATGATGATTCGGAATTCGCGCTTACGTC 1020
Db 5681 ATGGTTCGATCANTCGGCGCGATCGTGGATGATGATTCGGAATTCGCGCTTACGTC 5622
Qy 1021 TTCGGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGATGATC 1080
Db 5621 TTCGGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGATGATC 5562
Qy 1081 TTGCATATCGATTGCCGAGCGGATGTCGGCCCCCGGATTAATATCGTTTCGCAAGTGA 1140
Db 5561 TTGCATATCGATTGCCGAGCGGATGTCGGCCCCCGGATTAATATCGTTTCGCAAGTGA 5502
```

RESULT 5

ADH10182/c
ID ADH10182 standard; DNA; 11241 BP.

XX AC ADH10182;

XX AC

XX AC

XX AC

DT 11-MAR-2004 (first entry)

XX DE E. coli-Rhodococcus shuttle plasmid pRHB17.

XX DE Aryl carotenoid; beta-ionone; carotene desaturase; CrtU; isorenieratene;

XX DE chlorobactene; pharmaceutical; food supplement; animal feed additive;

XX DE food colourant; cosmetic; crtD; ds.

XX OS Synthetic.

XX OS

XX PN WO2003093200-A2.

XX

PD 13-NOV-2003.

XX PF 06-MAY-2003; 2003WO-US014111.

XX PR 06-MAY-2002; 2002US-0378312P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Cheng Q, Tao L, Rouviere PE;

XX WPI; 2004-081862/08.

XX PT Production of aryl carotenoids useful in e.g. pharmaceuticals and food

XX PT colorants involves bioconversion of cyclic carotenoids comprising beta-

XX PS ionone ring in the presence of carotene desaturase.

XX SQ Example 3; SEQ ID NO 13; 66pp; English.

CC The invention relates to production of aryl carotenoid compounds. The

CC method involves transforming a host cell comprising a cyclic carotenoid

CC having at least one beta-ionone ring with a gene encoding a carotene

CC desaturase (CrtU) followed by growing the transformed host cell. Also

CC provided is a method for regulating aryl carotenoid biosynthesis in a

CC host cell. The method is useful in the production of aryl carotenoids

CC such as isorenieratene and chlorobactene, which are useful in

CC pharmaceuticals, food supplements, animal feed additives, food colorants

CC and cosmetics. The present sequence represents a E. coli-Rhodococcus

XX shuttle plasmid pRHB17 nucleotide sequence

SQ Sequence 11241 BP; 2517 A; 3109 C; 2945 G; 2670 T; 0 U; 0 Other;

Query Match 100.0%; Score 1140; DB 12; Length 11241;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACCAGCTAAGTGTGAACACCTTTCGGGAAAAGACCGGCTCCCGTCTCTGTGTCG 60

Db 6641 ATGACCAGCTAAGTGTGAACACCTTTCGGGAAAAGACCGGCTCCCGTCTCTGTGTCG 6582

Qy 61 TCCGATAGCGCGCATCCGCGACGAACTCGCGACCCCAACTTCAACAAATCACCACGTC 120

Db 6581 TCCGATAGCGCGCATCCGCGACGAACTCGCGACCCCAACTTCAACAAATCACCACGTC 6522

Qy 121 GAAACATTTAACGCTGTGCGCGCGATTTCTGCGGTGAACCGGTGTGACCATTTGTCAAC 180

Db 6521 GAAACATTTAACGCTGTGCGCGCGATTTCTGCGGTGAACCGGTGTGACCATTTGTCAAC 6462

Qy 181 GGTCCGAAAGGTTCTGGATTCGGAGCGCTTCGTTCTCGGAAAAGGGCTGATCTGCCCC 240

Db 6461 GGTCCGAAAGGTTCTGGATTCGGAGCGCTTCGTTCTCGGAAAAGGGCTGATCTGCCCC 6402

Qy 241 TGCTGTGCGGAAAAGTTCGATGTCACATCTGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 300

Db 6401 TGCTGTGCGGAAAAGTTCGATGTCACATCTGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 6342

Qy 301 CAATCGGAGCTGATCTGTTGCGATGTCGACGATGACCATGCGCCATACAGCTGTCAG 360

Db 6341 CAATCGGAGCTGATCTGTTGCGATGTCGACGATGACCATGCGCCATACAGCTGTCAG 6282

Qy 361 CGGCTCCACGACCTATGAGTGGACTTTTCGGACGCTTGGAAAGCTGGACCAACGGTCTG 420

Db 6281 CGGCTCCACGACCTATGAGTGGACTTTTCGGACGCTTGGAAAGCTGGACCAACGGTCTG 6222

Qy 421 CGTTGGCGTACGGAAACGTTAGCGGTGCGACGATACGTCGCGCTGTTGAAATC 480

Db 6221 CGTTGGCGTACGGAAACGTTAGCGGTGCGACGATACGTCGCGCTGTTGAAATC 6162

Qy 481 ACTCAGGAAAAAAGCGGTGGCAGCTGCGACGTTTCAGCGCTACTCATGTTTCAGTGGTAC 540

Db 6161 ACTCAGGAAAAAAGCGGTGGCAGCTGCGACGTTTCAGCGCTACTCATGTTTCAGTGGTAC 6102

Qy 541 GTGAGTGAACATCTCTCGAATCCTTCTCGATGCGATGTTTCGATGCGGTGACATTCGAAA 600

```
Db 6101 GTGAGTGAGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTGACATTCCTCAA 6042
Qy 601 CTCGTATCTCTGGGATTTGCTGCGCACTAGCTAATTCGGTGGTCTCGATGTACGAAG 660
Db 6041 CTCGTATCTCTGGGATTTGCTGCGCACTAGCTAATTCGGTGGTCTCGATGTACGAAG 5982
Qy 661 ATCGGCGGTGAAGTCATCAAGTTCTCGTGGTATCTGACGAAAATTTGCATCTGGCGTT 720
Db 5981 ATCGGCGGTGAAGTCATCAAGTTCTCGTGGTATCTGACGAAAATTTGCATCTGGCGTT 5922
Qy 721 GGTATGAGGTTGGTGGGACGAGAAAGTGTGCACATGGCAACCGTGCACCTGG 780
Db 5921 GGTATGAGGTTGGTGGGACGAGAAAGTGTGCACATGGCAACCGTGCACCTGG 5862
Qy 781 GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAAGCGTTGGAATCTGTGGCGAATTT 840
Db 5861 GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAAGCGTTGGAATCTGTGGCGAATTT 5802
Qy 841 GAGTTGTTGATCGGAGCTCGGCAATCGGTGCTCGGTGATTTGGTCCCGAGCT 900
Db 5801 GAGTTGTTGATCGGAGCTCGGCAATCGGTGCTCGGTGATTTGGTCCCGAGCT 5742
Qy 901 GGTCTTGGGCGAAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGAAATCTGCCCGGTC 960
Db 5741 GGTCTTGGGCGAAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGAAATCTGCCCGGTC 5682
Qy 961 ATGTTTGGCATCATTTCCGGCGCATCGTGGATGATGATTCGGAATTTGTGGCGCTTACGTC 1020
Db 5681 ATGTTTGGCATCATTTCCGGCGCATCGTGGATGATGATTCGGAATTTGTGGCGCTTACGTC 5622
Qy 1021 TTCGGCGAGATCCTCGGACTGTCGAAGCTGGCGCATTCGGGAAAATCTTCGTGATCAC 1080
Db 5621 TTCGGCGAGATCCTCGGACTGTCGAAGCTGGCGCATTCGGGAAAATCTTCGTGATCAC 5562
Qy 1081 TTGCATTATCGATTGCCCGCAGCGGATGTCGGCCCCCGGATAATATCGGTTTCGCAAGTGA 1140
Db 5561 TTGCATTATCGATTGCCCGCAGCGGATGTCGGCCCCCGGATAATATCGGTTTCGCAAGTGA 5502
```

```
RESULT 6
AAV58945
ID AAV58945 standard; DNA; 2297 BP.
```

```
XX AC AAV58945;
```

```
XX DT 05-JAN-1999 (first entry)
```

```
XX DE B. breve essential region gene.
```

```
XX KW Essential region gene; shuttle vector; ds.
```

```
XX OS Bifidobacterium breve.
```

```
XX PH Key Location/Qualifiers
```

```
FT -35_signal 387..392
```

```
FT /tag= a
```

```
FT -10_signal 410..415
```

```
FT /tag= b
```

```
FT RBS 525..530
```

```
FT /tag= c
```

```
FT CDS 539..2023
```

```
FT /tag= d
```

```
XX JPI0262670-A.
```

```
XX PD 06-OCT-1998.
```

```
XX PF 27-MAR-1997; 97JP-00091387.
```

```
XX PR 27-MAR-1997; 97JP-00091387.
```

```
XX PA (HONS ) YAKULT HONSHA KK.
```

```
XX WPI; 1998-587288/50.
DR P-PSDB; AAW73071.
```

```
XX Shuttle vector for a Bifidobacterium species - contains specific elements from existing plasmids e.g. pNBb1 of ATCC 15698.
```

```
XX Disclosure; Page 14-15; 17pp; Japanese.
```

```
XX This sequence represents the essential region gene from plasmid pNBb1 used in the vector of the invention. The vector is a shuttle vector for a Bifidobacterium, and contains: (a) a replicated essential region CC originated from a plasmid pNBb1 of Bifidobacterium breve ATCC 15698; (b) a replicated essential region originated from a plasmid of E. coli; and CC (c) an antibiotic-resistant gene originated from a plasmid of E. coli and CC an antibiotic-resistant gene functioning by a Bifidobacterium. The CC shuttle vector can be used to transform various Bifidobacteria species
```

```
XX Sequence 2297 BP; 524 A; 659 C; 689 G; 425 T; 0 U; 0 Other;
```

```
Query Match 4.5%; Score 51.6; DB 2; Length 2297;
```

```
Best Local Similarity 52.9%; Fred. No. 3.5e-05;
```

```
Matches 111; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
```

```
Qy 394 GCCTGGAAGCTGCGACCAACGCTGCTGTTGGGCTACGGAACGTGAATGTACGGCTGC 453
```

```
Db 1184 GCTTGGACAGATGATTAAACGGAAGCCCTTGGCAACGGGCTCGGAACGTTGGAATC 1243
```

```
Qy 454 GACGGATACGTGCGCGCTGTTGAAATCACTCACGAAAAAACGGTGGCAGTCCACGTT 513
```

```
Db 1244 AGGGGTTTCGTCCGCGGATTGAAATCACCTACGTTGTAACGGCTGGCACCCTCACATT 1303
```

```
Qy 514 CACGGCTACTCATGTTTCAGTGTGACGTGAGTGAACATCCTCGAATCCTTCTCGGAT 573
```

```
Db 1304 CATTTGCGTCATGTTTCTCGATGGCGATCTGGACGATGGGCGTGGCAATGCAGCAA 1363
```

```
Qy 574 GCGATGTTTCGATCGGTGACATTCCAAATC 603
```

```
Db 1364 TGGCTGCTCGATCGCTGGAACCAATGTC 1393
```

```
RESULT 7
```

```
ADA71938
```

```
ID ADA71938 standard; DNA; 2000 BP.
```

```
XX AC ADA71938;
```

```
XX DT 20-NOV-2003 (first entry)
```

```
XX DE Rice gene, SEQ ID 5263.
```

```
XX KW Plant; bacterial infection; fungal infection; viral infection; rice; gene; ds.
```

```
XX OS Oryza sativa.
```

```
XX FN WO2003000898-A1.
```

```
XX PD 03-JAN-2003.
```

```
XX PF 22-JUN-2001; 2001WO-IB001105.
```

```
XX PR 22-JUN-2001; 2001WO-IB001105.
```

```
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
```

```
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
```

```
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
```

```
XX WPI; 2003-175290/17.
```

```
XX Identifying at least one gene involved in plant resistance or response to PT pathogenic infection for conferring resistance or tolerance to a plant to
```


KW vaccine; cancer; cytostatic; gene; ss.

XX Mus sp.
OS Synthetic.
XX Key Location/Qualifiers
FH CDS 72..1235
FT /*tag= a
FT /product= "modified PTX3"
PT
XX WO2003072603-A2.
XX
XX
XX PD 04-SEP-2003.
XX
XX PF 25-FEB-2003; 2003WO-IT000104.
XX
XX PR 28-FEB-2002; 2002IT-RM000109.
XX
XX PA (SIGT) SIGMA-TAU IND FARM RIUNITE SPA.
XX
XX PI De Santis R, Salvatori G;
XX
XX DR WPI; 2003-731602/69.
XX P-PSDB; ABG75068.
XX
XX PT New derivatives of human or murine pentraxin-3, useful for preparing
PT autologous vaccines for treating tumors, also new nucleic acid encoding
PT them.
XX
XX PS Claim 7; Page 31; 32pp; English.
XX
XX CC The present invention relates to derivatives of murine and human
CC pentraxin-3 (PTX3) (shown in ABG75068-ABG75073). These sequences, and
CC biotinylated derivatives of PTX3, are used to make autologous vaccines,
CC based on inactivated tumour cells, to treat solid or haematological
CC tumours. The present sequence is a derivative of the murine PTX3 cDNA
CC shown in the exemplification of the invention
XX
XX SQ Sequence 1244 BP; 261 A; 320 C; 405 G; 258 T; 0 U; 0 Other;

Query Match 3.6%; Score 40.6; DB 10; Length 1244;
Best Local Similarity 48.5%; Pred. No. 0.083;
Matches 112; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 696 TCTGACGAAATTGTCATCTGGCGTTGGTATGAGAGTTGGTAGTCGCGACGGAAGTGG 755
Db 331 TCCGTGGAGAGCTGCAGCGGCTGCGGCGAGAGCTGGGCGGCGCATGGCGA 390
QY 756 TCGACATGCGCAACCTGCACCTGGGAAATCGTCTTGATGAGTGGGCGGATCCACA 815
Db 391 GGCCGTGGCAGCCCGTGGCCCGCAGACGCGAGCTGGTGGGGCGCTGGAGCGCTGC 450
QY 816 AGCGTTGGAACCTGTGGCGAGAAATTGAGTTTGGTTTCGATGGGACGTCGCGGCAATCGCGTG 875
Db 451 TGCAGGAGAGCGCTGACGCGAGCCTCAGGCTGGCGCGCTGGAGACGCGGAGGCGCGC 510
QY 876 GTCCCGTGGATTGCGTCCCGAGCTGTTTGGGCGCAGACTAACAGATGC 926
Db 511 GACCCGAGCGCAGTGGCTGGCCCTAGGCGCTGTGCTGGAGGAATGCGGC 561

RESULT 10
ACC00046
ID ACC00046 standard; cDNA; 1841 BP.
XX
XX AC ACC00046;
XX
XX DT 11-AUG-2003 (first entry)
XX
XX DE Mouse PTX3 nucleotide sequence.
XX
XX KW Mouse pentraxin 3; PTX3; reproductive ability; antiinfertility;
KW infertility; ss.

XX Mus sp.

XX Key Location/Qualifiers
FH CDS 144..1286
FT /*tag= a
FT /product= "PTX3 mouse"
XX
XX WO2003011326-A1.
XX
XX PD 13-FEB-2003.
XX
XX PF 18-JUL-2002; 2002WO-IT000473.
XX
XX PR 03-AUG-2001; 2001US-0309472P.
XX
XX PA (SIGT) SIGMA-TAU IND FARM RIUNITE SPA.
XX
XX PI Mantovani A;
XX
XX DR WPI; 2003-239472/23.
XX P-PSDB; ABR58222.
XX
XX PT Use of recombinant human pentraxin 3 gene and protein for preparing a
PT medicament for increasing the reproductive ability in a female subject,
PT or for diagnosing or treating female infertility.
XX
XX PS Disclosure; Page 54; 63pp; English.
XX
XX CC The present invention relates to the use of recombinant human pentraxin 3
CC (PTX3) for preparing a medicament for increasing the reproductive ability
CC in a female subject, used as an antiinfertility drug. The virals or
CC plasmid vectors containing the human PTX3 cDNA are useful for treating
CC female subjects in need of increasing the reproductive ability. The PTX3
CC protein is useful as a diagnostic marker of the reproductive ability in
CC human female, and PTX3 is useful as a target protein for the screening of
CC pharmaceutical compounds to assess their capability to affect the
CC reproductive ability in a female subject. The present sequence represents
CC mouse PTX3 nucleotide sequence
XX
XX SQ Sequence 1841 BP; 444 A; 405 C; 527 G; 465 T; 0 U; 0 Other;

Query Match 3.6%; Score 40.6; DB 8; Length 1841;
Best Local Similarity 48.5%; Pred. No. 0.1;
Matches 112; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 696 TCTGACGAAATTGTCATCTGGCGTTGGTATGAGAGTTGGTAGTCGCGACGGAAGTGG 755
Db 341 TCCGTGGAGAGCTGCAGCGGCTGCGGCGAGAGCTGGGCGGCTGGCGCATGGCGA 400
QY 756 TCGACATGCGCAACCTGCACCTGGGAAATCGCTTTCATGAGTGGGCGGATCCACA 815
Db 401 GGCCGTGGCAGCCCGTGGCCCGCAGACGCGAGCTGGTGGGGCGCTGGAGCGCTGC 460
QY 816 AGCGTTGGAACCTGTGGCGAGAAATTGAGTTTGGTTTCGATGGGACGTCGCGGCAATCGCGTG 875
Db 461 TGCAGGAGAGCGCTGACGCGAGCCTCAGGCTGGCGCGCTGGAGACGCGGAGGCGCGC 520
QY 876 GTCCCGTGGATTGCGTCCCGAGCTGTTTGGGCGCAGACTAACAGATGC 926
Db 521 GACCCGAGCGCAGTGGCTGGCCCTAGGCGCTGTGCTGGAGGAATGCGGC 571

RESULT 11
ACC00048
ID ACC00048 standard; cDNA; 2708 BP.
XX
XX AC ACC00048;
XX
XX DT 11-AUG-2003 (first entry)
XX
XX DE Mouse upstream regulatory region.
XX

XX		Temperature sensitive plasmid; TSRCR; protein production;
KW		temperature sensitive replication control region; p48K; ds.
XX		Corynebacterium glutamicum.
OS		
XX	Key	Location/Qualifiers
FH		1318..2601
FT	CDS	/tag= a
FT		
XX	EPI039866-A1.	
PN		
XX	27-SEP-2000.	
XX	16-MAR-2000; 2000EP-00105326.	
XX	16-MAR-1999; 99JP-00069896.	
PR	(AJIN) AJINOMOTO CO INC.	
XX	Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;	
PI		
XX	WPI; 2000-573832/54.	
DR	P-ESDB; RAY97536.	
XX		Plasmids containing a temperature sensitive replication control regions useful for breeding microorganisms for the production of amino acids by fermentation.
PT		Claim 7; Page 14-17; 29pp; English.
PS		This sequence encodes the Brevibacterium lactofermentum p48K protein. The invention relates to a plasmid containing a temperature sensitive replication control region (TSRCR) and a marker gene (MG). The TSRCR is derived from plasmid pM330 harboured by Brevibacterium lactofermentum (ATCC 13869) and allows the plasmid to replicate autonomously at a low temperature but does not allow the plasmid to replicate autonomously at an elevated temperature in coryneform bacteria within a temperature range in which the bacteria can grow. The plasmid can be used for modifying a chromosomal gene in a coryneform bacterium, which may be used for the production of useful substances, such as amino acids, by fermentation, to change their genetic traits. Therefore, the plasmid can be used for breeding microorganisms for the production of amino acids by fermentation. The plasmid comprises a TSRCR that allows the plasmid to replicate autonomously at an elevated temperature in coryneform bacteria in a temperature range in which the bacteria can grow. (Updated on 15 SEP-2003 to standardise OS field)
XX		Sequence 4447 BP; 824 A; 997 C; 1378 G; 1248 T; 0 U; 0 Other;
SQ		
	Query Match	3.3%; Score 37.4; DB 3; Length 4447;
	Best Local Similarity	53.0%; Pred. No. 1.7;
	Matches	80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY	493 AACGGCTGGCAGTCCACGTTCAACCGGTACTCATGTTTCAGTGAGTGAGAAC	552
Db	1795 AACGGTTGGCACTGCACGCCAACAATGCTGTGTTCTTGGATCGTCCACTGCTGACGAT	1854
QY	553 ATCCTCGAATCTTCTCGATGCGATGTTTCGATCGGTGGACCTTCCAACACTGTAATCTCTG	612
Db	1855 GAACCTCAAGGGCTTTGAGGAATCCATGTTTTCCCCTGCTGCTGTTGTTAAAGCC	1914
QY	613 GGATTTCGTCGCCCACTACGTAATTCGGGTG	643
Db	1915 GGTATGGACGCCCACTGCGTGAGCACGGGG	1945
	RESULT 14	
ABU49734		
ID	ABL49734 standard; DNA; 4447 BP.	
XX	AC	ABL49734;
XX		

DT	29-AUG-2003 (revised)	
DT	29-MAY-2002 (first entry)	
XX	Brevibacterium lactofermentum p48K nucleotide sequence SEQ ID NO:7.	
DE		
XX	Kanamycin resistant gene; mutation; high temperature resistance;	
KW	coryneform microbe; transformation; gene; ds.	
XX		
OS	Corynebacterium glutamicum.	
XX		
FH	Key Location/Qualifiers	
FT	1318..2601	
FT	*tag= a	
FT	/product= "p48K"	
XX	JP2002017362-A.	
PN		
XX	22-JAN-2002.	
PD		
XX	05-JUL-2000; 2000JP-00204236.	
PF		
XX	05-JUL-2000; 2000JP-00204236.	
PR		
XX	(AJIN) AJINOMOTO KK.	
PA		
XX	WPI; 2002-263243/31.	
DR	P-PSDB; ABR06342.	
XX	Transformation of high temperature-resistant coryneform microbe and transformant.	
PT		
PT	Example; Page 20-22; 28pp; Japanese.	
PS		
XX	The present invention describes a method for transforming a high temperature-resistant coryneform microbe in which the high temperature-resistant coryneform microbe is treated with a chemical changing the structure of cell wall of the microbe and an electric pulse is applied to a solution containing the microbe having a changed structure of cell wall and a DNA. Also described are: (1) transforming a high temperature-resistant coryneform microbe in which the cell wall of a high temperature-resistant coryneform microbe is weakened and electric pulse is applied to a solution containing the microbe having weakened cell wall and a DNA; (2) a recombinant DNA is introduced; and (3) a transformant of a high temperature-resistant coryneform microbe carrying both a plasmid vector derived from the high temperature-resistant coryneform microbe and a plasmid vector derived from Corynebacterium glutamicum or Brevibacterium lactofermentum. The method is used for transforming a high temperature-resistant coryneform microbe. The present sequence encodes a Brevibacterium lactofermentum p48K protein sequence, which is used in an example from the present invention. (Updated on 29-AUG-2003 to standardise OS field)	
XX		Sequence 4447 BP; 827 A; 993 C; 1375 G; 1252 T; 0 U; 0 Other;
SQ		
	Query Match	3.3%; Score 37.4; DB 6; Length 4447;
	Best Local Similarity	53.0%; Pred. No. 1.7;
	Matches	80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY	493 AACGGTGGCAGTCCACGTTCAACCGGTACTCATGTTTCAGTGAGTGAGAAC	552
Db	1795 AACGGTTGGCACTGCACGCCAACAATGCTGTGTTCTTGGATCGTCCACTGCTGACGAT	1854
QY	553 ATCCTCGAATCTTCTCGATGCGATGTTTCGATCGGTGGACCTTCCAACACTCGTATCTCTG	612
Db	1855 GAACCTCAAGGCATTTGAGGATTCATGTTTTCCCGCTGCTGCTGTTGTTAAAGCC	1914
QY	613 GGATTTCGTCGCCCACTACGTAATTCGGGTG	643
Db	1915 GGTATGGACGCCCACTGCGTGAGCACGGGG	1945
	RESULT 15	

	Query Match	3.1%	Score 34.8	DB 4	Length 1631	
	Best Local Similarity	50.0%	Fed. No. 0.94			
	Matches 87	Conservative	0	Mismatches 87	Indels 0	Gaps 0
Qy	219	CGGAAAGGGCTGGATCTGCCCTCTGTGCGGGAAAAGTCGTGCACATCGTGCAGACGA	278			
Db	1184	CGGAAATGTCCTGGAGCGACCGAGGATGTCAAAAAAAATGGCGCGAAACGTGATCTGCCGC	1243			
	279	AATTTCTCAAGCTGTTGCTCATCACTCGGCACCTGGATCTCTTGGCATGGTGACGATGAC	338			

	Query Match	3.0%	Score 34.6;	DB 1;	Length 1505;
	Best Local Similarity	58.1%;	Pred. No. 1;		
	Matches 61;	Conservative 0;	Mismatches 44;	Indels 0;	Gaps 0;
Qy	706	ATTGCATCTCGCGTTGCTATGAGGTTGGTAGTGGCGACGAAAAAAGTGTGCACATGCC	765		
Db	699	ATTGGAGTTGATGCTGCTGTTTGGTTTGGAGGTGTCGCGGAGAGAGGTGGTCTGGT	758		
Qy	766	AACCGTGCACCCCTGGGAAATCCGCTGTTGATGTCAGTGGGCGCGGGAT	810		
Db	759	TACGTTGAGCTGAAGCACTTGGTTGGAGGAGGAGCGCGTGGACAT	803		

338

```
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match          3.0%; Score 34.4; DB 1; Length 7218;
Best Local Similarity 2.4%; Pred. No. 2.9;
Matches 8; Conservative 186; Mismatches 142; Indels 0; Gaps 0;

Qy 613 GGATTGTCGGCCACTACGTAATTCGGGTGGTCTCGATGTACGAAAGATCGCGGTGAA 672
| ||||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1445 GAATTGTCACRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1386

Qy 673 GCTGATCAGTCTCGCTGCGTATCTGACGAAATTCGATCTGGCGTTGTGAGAGTT 732
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1385 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1326

Qy 733 GGTAGTCGCGACGAAAGGTGTCGACATGCGAACCGTCGACCTGGGAAATCGCTGT 792
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1325 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1266

Qy 793 GATCAGTGGCGGGGATCCACAGCTGTGGAACGTGGCGAGAAATTTGAGTTGGTTCG 852
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1265 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1206

Qy 853 ATGGACCTCGGCAATCGCTGGTCCCGTGGATTCGTCGCGAGCTGCTTTGGGGCA 912
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1205 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1146

Qy 913 GAACCTAACAGTCTCAGATCGTTGACGACGAGNA 948
```

```
Db 1145 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1110

RESULT 6
US-09-482-180A-3/c
; Sequence 3, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; FILE REFERENCE: BETA-1.3-GALACTOSYLTRANSFERASE HOMOLOG, ZN5SP6
; CURRENT APPLICATION NUMBER: US/09/482,180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: variation
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: misc_feature
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n = A,T,C or G
US-09-482-180A-3

Query Match          2.9%; Score 33.6; DB 3; Length 1134;
Best Local Similarity 30.3%; Pred. No. 1.9;
Matches 61; Conservative 30; Mismatches 110; Indels 0; Gaps 0;

Qy 417 TCGTCTGTGGCTACGGAACGTGAAATGTACGGCTGGCAGCATACGTGCGCTGTGA 476
| : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 796 TNSWNGGNGDATRAARTAYTTNACYTTNGTTRTNCRTTNGNARNGCYTGNCCKDNA 737

Qy 477 AATCACTCACGGAAGAAACGGCTGGCAGCTTCACGCGTACTCATGTTTCAGTGG 536
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 736 CRTCCNACNARNARNTCYTNGCNGGRTCCCAACCCRTCNARRAAYTCNARNACRTTG 677

Qy 537 TGAGCTGAGTGAGAACATCTCGAATCTTCTCGATGCGATGTTGATCGGTGACTTC 596
| : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 676 GNACRTGNACRAANACRTCTCTCNCCTTTCNARCAATRAARTNGCYTNGGRCANGCNG 617

Qy 597 CAAACTCGTATCTCTGGGATT 617
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 616 CNACNACCAACKYTGNAART 596

RESULT 7
US-08-813-150-3
; Sequence 3, Application US/08813150
; Patent No. 6069229
; GENERAL INFORMATION:
; APPLICANT: Mueller, Christopher
; APPLICANT: Lebecque, Serge J.E.
; APPLICANT: Liu Yong-Jun
; APPLICANT: Dowling, Lynette M.
; APPLICANT: Huffine, Constance M.
; APPLICANT: Gorman, Daniel M.
; TITLE OF INVENTION: MAMMALIAN PROTEINASES; OXIDOREDUCTASES;
; TITLE OF INVENTION: RELATED REAGENTS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/813,150
FILING DATE: 07-MAR-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0693
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-852-9196
TELEFAX: 650-496-1200
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1780 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 66..1766
US-08-813-150-3

Query Match 2.9%; Score 33.2; DB 3; Length 1780;
Best Local Similarity 55.1%; Pred. No. 3.2;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 640 GGTGCTCTCGATGACGAAGATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 699
DB 1016 GGTGCTCTCGATGACGAAGATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 1075
QY 700 ACGAAATTTGATCTGGCGTTGGTATGAGGCTGAGGAGGCTGACACTACGTGCGGCGCCACCAAGTGTTC 757
DB 1076 CCGCCACATGACGAGGAGGCTGAGGAGGCTGACACTACGTGCGGCGCCACCAAGTGTTC 1133

RESULT 8
US-09-546-553-3
Sequence 3, Application US/09546553
Patent No. 6518405
GENERAL INFORMATION:
APPLICANT: Mueller, Christopher
APPLICANT: Lebecque, Serge J.E.
APPLICANT: Liu, Yong-Jun
APPLICANT: Dowling, Lynette M.
APPLICANT: Huffine, Constance M.
APPLICANT: Gorman, Daniel M.
TITLE OF INVENTION: MAMMALIAN PROTEINASES; OXIDOREDUCTASES;
TITLE OF INVENTION: RELATED REAGENTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/546,553
FILING DATE: 10-APR-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/813,150
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0693
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-852-9196
TELEFAX: 650-496-1200
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1780 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 66..1766
US-09-546-553-3

Query Match 2.9%; Score 33.2; DB 4; Length 1780;
Best Local Similarity 55.1%; Pred. No. 3.2;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 640 GGTGCTCTCGATGACGAAGATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 699
DB 1016 GGTGCTCTCGATGACGAAGATCGCGGTGAAGCTGATCAAGTTCTCGCTCGGTATCTG 1075
QY 700 ACGAAATTTGATCTGGCGTTGGTATGAGGCTGAGGAGGCTGACACTACGTGCGGCGCCACCAAGTGTTC 757
DB 1076 CCGCCACATGACGAGGAGGCTGAGGAGGCTGACACTACGTGCGGCGCCACCAAGTGTTC 1133

RESULT 9
US-10-140-002-475
Sequence 475, Application US/10140002
Patent No. 6725730
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RIC59
CURRENT APPLICATION NUMBER: US/10/140,002
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 475
LENGTH: 1844
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-002-475

Query Match 2.9%; Score 33.2; DB 4; Length 1844;
Best Local Similarity 55.1%; Pred. No. 3.3;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 640 GGTGCTCTCGATGACGAAAGATCGCGGTGAAGTGTATCAAGTCTCGCTGCGTATCG 699
DB 1029 GGTGCTGCTGACGCGGAGCGGACCGCGGTGAAGCGCATCACTTCTCGCGCGCTGCC 1088
QY 700 ACGAAATTCATCTGCGTGTGGTATGAGGTTGGTGTGCGGACGGAAGTGGTC 757
DB 1089 CCGCCACATGACGAGGGGCTCGGAGGCTGCACTACGTGCGCGCCACCAAGGTGTC 1146

RESULT 10

US-09-482-273-81
; Sequence 81, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 1867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-81

Query Match 2.9%; Score 33.2; DB 4; Length 1867;
Best Local Similarity 55.1%; Pred. No. 3.3;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 640 GGTGCTCTCGATGACGAAAGATCGCGGTGAAGTGTATCAAGTCTCGCTGCGTATCG 699
DB 1053 GGTGCTGCTGACGCGGAGCGGACCGCGGTGAAGCGCATCACTTCTCGCGCGCTGCC 1112
QY 700 ACGAAATTCATCTGCGTGTGGTATGAGGTTGGTGTGCGGACGGAAGTGGTC 757
DB 1113 CCGCCACATGACGAGGGGCTCGGAGGCTGCACTACGTGCGCGCCACCAAGGTGTC 1170

RESULT 11

US-08-722-001-26
; Sequence 26, Application US/08722001
; Patent No. 5760054
; GENERAL INFORMATION:
; APPLICANT: Thompson, Wayne J.
; APPLICANT: Huff, Joel R.
; APPLICANT: Nerenberg, Jennie B.
; APPLICANT: Lee, Hee-Yoon
; APPLICANT: Bell, Ian M.
; TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07065
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,001
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,276
FILING DATE: 14-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Appollina, Mary A.
REGISTRATION NUMBER: 34,087
REFERENCE/DOCKET NUMBER: 19169Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)594-3462
TELEFAX: (908)594-4720
TELEX: 138825

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1987 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-722-001-26

Query Match 2.9%; Score 32.8; DB 1; Length 1987;
Best Local Similarity 56.5%; Pred. No. 4.6;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 105 ACAATACCAAGTCAGAAACATTTACGCTGTGCGCGCGCGATTCTTGGCGTGAACGG 164
DB 778 ACGCACTACTACATCGTCAACCGCTAGTGGCGGTGGCGACCTCTGCTCACCTCCACGG 837
QY 165 TGTGACCATTTGCAACGGTCCGAAAGTTCTTGATTCCGGAGGCTTCG 212
DB 838 TGCTGCTCTCTCCGCCATCTTCGAGGTCCTAGGCTACTGGGCTTCG 885

RESULT 12

US-08-209-747-1
; Sequence 1, Application US/08209747
; Patent No. 5733771
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Colgin, Mark
; TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,747
FILING DATE: 14-MAR-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P

Db 209 RKSSKKRYATYYSCMKWKYCMMSATYSGMMRWYCYSCMMRYSTYSRGKCS 150
QY 858 ACGTCCGGCAATCCGTCGTCCTCCGTCGATTCGTCGCCGAGCTGCTCTTGGGGCAGAACT 917
Db 149 CTGWGKCYSYRMVGRWKYRMSRGARRYTKSRGWCRCSTKRYKTCASWGAAGTMC 90
QY 918 AACAGATGCTCAGATCGTTGAGCAGGAAGAATCTGCCCGGTCATGTTGGATCATCC 977
Db 89 CMRGTGAGTMYRMKSYKRYKKWSWKSMKGYGYSWSYMSGSGCCKSCGSSSYCRSYSS 30
QY 978 GGCGCGATCGTGATGATCGGACTT 1006
Db 29 MGKSRCTCCMGCGGYTYCMKCSRYT 1

RESULT 15
US-09-614-912-49/c
; Sequence 49, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Cai, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; PENDING FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 49
; TYPE: DNA
; LENGTH: 1431
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1309)
; NAME/KEY: unsure
; LOCATION: (1339)
; NAME/KEY: unsure
; LOCATION: (1351)
; NAME/KEY: unsure
; LOCATION: (1402)
; NAME/KEY: unsure
; LOCATION: (1429)
US-09-614-912-49

Query Match 2.9%; Score 32.6; DB 4; Length 1431;
Best Local Similarity 49.7%; Pred. No. 4.5;
Matches 83; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 890 GTGCCCGAGTCGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAT 949
Db 850 GCGCCCGCGCGGCGCTGATGAGGCGCACCGAGCGTGGCGCGCTCGAACACCTGGAGT 791

QY 950 CTGCCCGGTCAATGTTGCGATCAATCCGGCGCATCGTGGATGATGATTCGGAATTG 1009
Db 790 CCGCCACAGTCGTCCTGTCGAGATGGTGAGCGACGTCGGGTGCGCAATCGGCCCGGTG 731
QY 1010 CGCCTTACGTCTTCCGGCGAGATCCTCGGACTCTGTCGAAGCTGGGGCG 1056
Db 730 CCCAGCGTGTGCTAGGCGCGCTGGCACGCGGGGTAGTAGTTGAGGCG 684

Search completed: October 23, 2004, 02:20:26
Job time : 129 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2004, 01:10:16 ; Search time 618 Seconds

(without alignments)
9451.893 Million cell updates/sec

Title: US-10-007-527A-1

Perfect score: 1140

Sequence: 1 atgaccagcgaagtgtgta.....taatatcggttcgaagtga 1140

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3407233 seqs, 2561960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	100.0	1140	14	US-10-007-527A-1
2	1140	100.0	1140	14	US-10-007-452-1
3	1140	100.0	1140	17	US-10-415-562A-1
4	1140	100.0	6334	14	US-10-007-527A-5
5	1140	100.0	6334	14	US-10-007-452-5
6	1140	100.0	6334	17	US-10-415-562A-5
7	1140	100.0	9652	14	US-10-007-527A-7
8	1140	100.0	9652	14	US-10-007-452-7
9	1140	100.0	9652	17	US-10-415-562A-7
10	1140	100.0	11241	14	US-10-007-527A-6
11	1140	100.0	11241	14	US-10-007-452-6
12	1140	100.0	11241	17	US-10-415-562A-6

13	40.8	3.6	520	14	US-10-184-644-332	Sequence 332, App
14	40.8	3.6	520	14	US-10-184-634-332	Sequence 332, App
15	1841	3.6	1841	18	US-10-485-640-3	Sequence 3, Appli
16	38.2	3.4	3794	16	US-10-424-599-41971	Sequence 41971, A
17	37.6	3.3	594	14	US-10-123-155-10	Sequence 10, Appl
18	37.6	3.3	594	15	US-10-146-731-10	Sequence 10, Appl
19	37.6	3.3	594	15	US-10-140-472-10	Sequence 10, Appl
20	37.6	3.3	594	15	US-10-141-761-10	Sequence 10, Appl
21	37.6	3.3	594	15	US-10-142-885-10	Sequence 10, Appl
22	37.6	3.3	594	15	US-10-158-790-10	Sequence 10, Appl
23	37.6	3.3	594	15	US-10-137-871-10	Sequence 10, Appl
24	37.6	3.3	594	15	US-10-140-923-10	Sequence 10, Appl
25	37.6	3.3	594	15	US-10-141-756-10	Sequence 10, Appl
26	37.6	3.3	594	15	US-10-141-759-10	Sequence 10, Appl
27	37.6	3.3	594	15	US-10-140-805-10	Sequence 10, Appl
28	37.6	3.3	594	15	US-10-140-864-10	Sequence 10, Appl
29	37.6	3.3	594	16	US-10-142-426-10	Sequence 10, Appl
30	37.4	3.3	2708	18	US-10-485-640-6	Sequence 6, Appli
31	37.4	3.3	4447	9	US-09-835-381-5	Sequence 5, Appli
32	37.4	3.3	4447	9	US-09-835-381-7	Sequence 7, Appli
33	37.4	3.3	4447	15	US-10-196-232-5	Sequence 5, Appli
34	37.4	3.3	4447	15	US-10-196-232-6	Sequence 6, Appli
35	36.6	3.2	1675	16	US-10-425-114-15202	Sequence 15202, A
36	36	3.2	299	14	US-10-184-644-164	Sequence 164, App
37	36	3.2	299	14	US-10-184-634-164	Sequence 164, App
38	36	3.2	1193	17	US-10-437-963-46791	Sequence 46791, A
39	36	3.2	1560	17	US-10-437-963-6403	Sequence 6403, Ap
40	36	3.2	2365	16	US-10-424-599-125091	Sequence 125091,
41	34.8	3.1	498	16	US-10-282-122A-6814	Sequence 6814, Ap
42	34.8	3.1	702	16	US-10-282-122A-1882	Sequence 1882, Ap
43	34.8	3.1	861	16	US-10-282-122A-41252	Sequence 41252, A
44	34.8	3.1	1631	9	US-09-912-020-208	Sequence 208, App
45	34.4	3.0	999	14	US-10-184-644-434	Sequence 434, App

ALIGNMENTS

RESULT 1
US-10-007-527A-1
; Sequence 1, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kisty N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CIL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-007-527A-1

Query Match 100.0%; Score 1140; DB 14; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACCAGCGTAACTGCTGACACCTTTCGGGCAAGACCGGCTCCGCTCTCGTCTCG 60
Db 1 ATGACCAGCGTAACTGCTGACACCTTTCGGGCAAGACCGGCTCCGCTCTCGTCTCG 60

Qy 61 TCCGATTAAGCGCGCATCCGGCAGCAACTCGACCCCAAACTTCAACAAATCACCACGTCA 120
Db 61 TCCGATTAAGCGCGCATCCGGCAGCAACTCGACCCCAAACTTCAACAAATCACCACGTCA 120


```
Db 721 GGTATGGAGTTGGTAGTGGCGACGGAAGAGTGGTCGACATGGCAACCGTCGACCCCTGG 780
Qy 781 GAAATCGCTGTTGATGCTGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Db 781 GAAATCGCTGTTGATGCTGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Qy 841 GAGTTTGGTTCCATGGGAGTGGGCAATCGGCTGCTCCGCTGGATGCTGCCGAGCT 900
Db 841 GAGTTTGGTTCCATGGGAGTGGGCAATCGGCTGCTCCGCTGGATGCTGCCGAGCT 900
Qy 901 GGTCTGGGCGAGAACTAACAGATCTCAGATCGTTGACGAGAGAAATCTGCCCGGTC 960
Db 901 GGTCTGGGCGAGAACTAACAGATCTCAGATCGTTGACGAGAGAAATCTGCCCGGTC 960
Qy 961 ATGGTGGCGATCATTCGCGGCGAGTCGGTGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Db 961 ATGGTGGCGATCATTCGCGGCGAGTCGGTGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Qy 1021 TTCGGCGAGATCTCGGACTCGTCGAGCTGGCGGAGCTTGGGAAATCTTTCGTGATCAC 1080
Db 1021 TTCGGCGAGATCTCGGACTCGTCGAGCTGGCGGAGCTTGGGAAATCTTTCGTGATCAC 1080
Qy 1081 TTGCATTATCGATTGCCCGAGCGGATGTGGCGCCCGGATAATTCGGTTCCGCAAGTGA 1140
Db 1081 TTGCATTATCGATTGCCCGAGCGGATGTGGCGCCCGGATAATTCGGTTCCGCAAGTGA 1140

RESULT 3
US-10-415-562A-1
; Sequence 1, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: Cl1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; PRIOR FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-415-562A-1

Query Match
Best Local Similarity 100.0%; Score 1140; DB 17; Length 1140;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACCAGGTAAGTGTGAACACCTTTCCGCAAGACCGGCTCCCGTCTCTCGTCTCG 60
Db 1 ATGACCAGGTAAGTGTGAACACCTTTCCGCAAGACCGGCTCCCGTCTCTCGTCTCG 60
Qy 61 TCCGATAGCGCGGATCCGCGACGAACTGGACCCAACTTCAACAAATCACCACGTCA 120
Db 61 TCCGATAGCGCGGATCCGCGACGAACTGGACCCAACTTCAACAAATCACCACGTCA 120
Qy 121 GAAACATTTAAACGCTGTGGCGGCGGATTTCTGGCGTGAACGCTGTGACCATTTGCAAC 180
Db 121 GAAACATTTAAACGCTGTGGCGGCGGATTTCTGGCGTGAACGCTGTGACCATTTGCAAC 180
Qy 181 GGTCCGAAAGTTCTGATTCGAGGCTTCTGTTCTCGGAAAGGCTGGATCTGCC 240
Db 181 GGTCCGAAAGTTCTGATTCGAGGCTTCTGTTCTCGGAAAGGCTGGATCTGCC 240
Qy 241 TGCTGTCGGGAAAGTCTGATTCGATCTGTCGACAGAAATTTCTCAAGTTGTGCTCAT 300
Db 241 TGCTGTCGGGAAAGTCTGATTCGATCTGTCGACAGAAATTTCTCAAGTTGTGCTCAT 300
Qy 301 CAACTCGGAGTCTGATCTGTTGGATGGTACCATGACCATGCGCCATACAGTGTGTCAG 360
Db 301 CAACTCGGAGTCTGATCTGTTGGATGGTACCATGACCATGCGCCATACAGTGTGTCAG 360
```

```
Db 301 CAACTCGGAGTCTGATCTGTTGCGATGTGACGATGACCATGCCATACAGTGTGTCAG 360
Qy 361 CGGCTCCACGACTATGAGTGGACTTTTCGGAGCGCTGGAAAGCTCGGACCAACGCTCGT 420
Db 361 CGGCTCCACGACTATGAGTGGACTTTTCGGAGCGCTGGAAAGCTCGGACCAACGCTCGT 420
Qy 421 CTTTGGCGTACGAAACGTTGAAATGTACCGCTCGGACGATACGTCGGCGCTGTTGAAATC 480
Db 421 CTTTGGCGTACGAAACGTTGAAATGTACCGCTCGGACGATACGTCGGCGCTGTTGAAATC 480
Qy 481 ACTCAGGAAAGAAACGCTGCGACGCTCCACGTTTACGCGCTACTCATGTTAGTGTGAC 540
Db 481 ACTCAGGAAAGAAACGCTGCGACGCTCCACGTTTACGCGCTACTCATGTTAGTGTGAC 540
Qy 541 GTGAGTGAGAAATCTCTCGAATCTTTCGAAATCTTCGATGCGATGTTTCGATGCGGCTTCCAAA 600
Db 541 GTGAGTGAGAAATCTCTCGAATCTTTCGAAATCTTCGATGCGATGTTTCGATGCGGCTTCCAAA 600
Qy 601 CTGATCTCTGGGATTTGTCGCGCACTACGTAATTCGGTGTCTCGATGTACGAAAG 660
Db 601 CTGATCTCTGGGATTTGTCGCGCACTACGTAATTCGGTGTCTCGATGTACGAAAG 660
Qy 661 ATCGCGGTGAGCTGATCAAGTTCTCGTGCCTCTGACGAAATTCGATGCGGCTTCCAAA 720
Db 661 ATCGCGGTGAGCTGATCAAGTTCTCGTGCCTCTGACGAAATTCGATGCGGCTTCCAAA 720
Qy 721 GGTATGGAGGTTGGTAGTGGCGGAGGAAAGTGGTCGACATGGCAACCGTGCACCCCTGG 780
Db 721 GGTATGGAGGTTGGTAGTGGCGGAGGAAAGTGGTCGACATGGCAACCGTGCACCCCTGG 780
Qy 781 GAAATCGCTGTTGATGCTGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Db 781 GAAATCGCTGTTGATGCTGAGTGGCGGGGATCCACAAGCGTTGGAATCTGTGGCGAGAAATTT 840
Qy 841 GAGTTTGGTTCGATGGGCAATCGCGCAATCGCGTGGTCCCGTGGATGCTGCCGAGCT 900
Db 841 GAGTTTGGTTCGATGGGCAATCGCGCAATCGCGTGGTCCCGTGGATGCTGCCGAGCT 900
Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATGATGATTCGATGCGGCTTACGTC 960
Db 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATGATGATTCGATGCGGCTTACGTC 960
Qy 961 ATGGTGGCGATCATTCGCGGCGAGTCGGTGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Db 961 ATGGTGGCGATCATTCGCGGCGAGTCGGTGATGATGATTCGGACTTGTGGCGCTTACGTC 1020
Qy 1021 TTCGGCGAGATCTCGGACTCGTCGAGCTGGCGGAGCTTGGGAAATCTTTCGTGATCAC 1080
Db 1021 TTCGGCGAGATCTCGGACTCGTCGAGCTGGCGGAGCTTGGGAAATCTTTCGTGATCAC 1080
Qy 1081 TTGCATTATCGATTGCCCGAGCGGATGTGGCGCCCGGATAATTCGGTTCCGCAAGTGA 1140
Db 1081 TTGCATTATCGATTGCCCGAGCGGATGTGGCGCCCGGATAATTCGGTTCCGCAAGTGA 1140

RESULT 4
US-10-007-527A-5/c
; Sequence 5, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: Cl1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
```

QY	961	ATGGTTGC	GATCAT	TCCGGC	GCATCG	TGCGAT	GATGAT	TCGCGCTT	ACGTC	1020	
Db	2091	ATGGTTGC	GATCAT	TCCGGC	GCATCG	TGCGAT	GATGAT	TCGCGCTT	ACGTC	2032	
QY	1021	TTCCGCGA	GATCCT	CGGACT	CTGCGA	AGCTGG	CGCGACT	TTGGAA	NAATCTT	CGTGATCAC	1080
Db	2031	TTCCGCGA	GATCCT	CGGACT	CTGCGA	AGCTGG	CGCGACT	TTGGAA	NAATCTT	CGTGATCAC	172
QY	1081	TTGCATTAT	CGATTG	CCCGAG	CGGATG	TCGGCC	CCCCCG	CAATAATAT	TCGGTTC	CAAGTGA	1140
Db	1971	TTGCATTAT	CGATTG	CCCGAG	CGGATG	TCGGCC	CCCCCG	CAATAATAT	TCGGTTC	CAAGTGA	1912

RESULT 5

```

US-10-007-452-5/C
; Sequence 5, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: Cbl1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
; US-10-007-452-5

```

Query Match	100.0%;	Score 1140;	DB 14;	Length 6334;
Best Local Similarity	100.0%;	Pred. No. 0;	Indels 0;	Gaps 0

1	QY	ATGACCAGCTAAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCGGCTCTCGTGTGG	299
3051	Db	ATGACCAGCTAAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCGGCTCTCGTGTGG	299
61	QY	TCCGATAGCGCGCATCCGGCAGCAACTCGGACCCCAACTTCAACAATACCAAGTCA	120
2991	Db	TCCGATAGCGCGCATCCGGCAGCAACTCGGACCCCAACTTCAACAATACCAAGTCA	293
121	QY	GAACAATTAAAGCTGTGGCGCGCGGATTTCTGGCGTGAACGGTGTGACATTGTCAAC	180
2931	Db	GAACAATTAAAGCTGTGGCGCGCGGATTTCTGGCGTGAACGGTGTGACATTGTCAAC	28
181	QY	GGTCCGAAGGTTCTGGATTCCGAGGCGTTCGTCTTCGGGAAGGCTGGATCTGCCCC	240
2871	Db	GGTCCGAAGGTTCTGGATTCCGAGGCGTTCGTCTTCGGGAAGGCTGGATCTGCCCC	281
241	QY	TGCTGTGCGGAAAAAGTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT	300
2811	Db	TGCTGTGCGGAAAAAGTCGGTGCACATCGTGCAGACGAAATTTCTCAAGTTGTTGCTCAT	275
301	QY	CAACTCGGACTTGGATCTGTCGGATGTGTGACATGACCATCGGCCATACAGCTGGTCAG	361
2751	Db	CAACTCGGACTTGGATCTGTCGGATGTGTGACATGACCATCGGCCATACAGCTGGTCAG	261
361	QY	CGGCTCCACGACTATGACTCGACTTTCCGACAGCTCGAAAGCTCGCACCAACGGTGGT	421
2691	Db	CGGCTCCACGACTATGACTCGACTTTCCGACAGCTCGAAAGCTCGCACCAACGGTGGT	261
421	QY	CGTGGCGTACGAAAGTGAATGTACGGCTCGACGGATACGTCGGCGCTGTTGAATC	481
2631	Db	CGTGGCGTACGAAAGTGAATGTACGGCTCGACGGATACGTCGGCGCTGTTGAATC	251
481	QY	ACTCACGGAAAAAACGGCTGGCATGTCACGGTTCACGGCTACTCATGTTTCAGTGGTGAC	541

; LENGTH: 6334
 ; TYPE: DNA
 ; ORGANISM: Rh
 US-10-007-527A-5

Query Match	100.0%;	Score: 11140;	DB 14;	Length 6334;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1140;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	ATGACACAGCGTAAGTGTCTGAACACCTTTTCGGCGAAAGACCGGCGCTCCCGTCTCGTGTGCG	60		
3051	ATGACACAGCGTAAGTGTCTGAACACCTTTTCGGCGAAAGACCGGCGCTCCCGTCTCGTGTGCG	2992		
61	TCCGATTAAGCGCGGCGATCCGGCAGCAAGCTCGACCCAAATCTCAACAAATCACCACGTC	120		
2991	TCCGATTAAGCGCGGCGATCCGGCAGCAAGCTCGACCCAAATCTCAACAAATCACCACGTC	2932		
121	GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCAATCTCAAC	180		
2931	GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCAATCTCAAC	2872		
181	GGTCGGAAGGTTCTGATTTCCGAGGCTTCGTTCTTCGGAAAGGCGTGGATCTGCCCC	240		
2871	GGTCGGAAGGTTCTGATTTCCGAGGCTTCGTTCTTCGGAAAGGCGTGGATCTGCCCC	2812		
241	TGCTGTCCGGGAAAGTCGGTGCACATCTGTCGACAGAAATTTCTCAAGTTGTTGCTCAT	300		
2811	TGCTGTCCGGGAAAGTCGGTGCACATCTGTCGACAGAAATTTCTCAAGTTGTTGCTCAT	2752		
301	CAACTCGGGAATGATCTGTTGCGATGTCAGCATGACCATCGCCCATACAGCTGGTCTAG	360		
2751	CAACTCGGGAATGATCTGTTGCGATGTCAGCATGACCATCGCCCATACAGCTGGTCTAG	2692		
361	CGGCTCCACGACCTATGGAATGGAATTTTCGGCAGCTTGGAAAGCTGCGACCAAGTCGT	420		
2691	CGGCTCCACGACCTATGGAATGGAATTTTCGGCAGCTTGGAAAGCTGCGACCAAGTCGT	2632		
421	CGTTGGCGGTACGGAACGTTGAAATGTACGGCTTCGCAAGCATACGTCGCGCTGTTGAAATC	480		
2631	CGTTGGCGGTACGGAACGTTGAAATGTACGGCTTCGCAAGCATACGTCGCGCTGTTGAAATC	2572		
481	ACTCACGGAAAAAAGCGGTGGCAGCTCCACGTTTCAACGGCTTCACTGTTCACTGTTGAGTGC	540		
2571	ACTCACGGAAAAAAGCGGTGGCAGCTCCACGTTTCAACGGCTTCACTGTTCACTGTTGAGTGC	2512		
541	GTGAGTGAGAACATCTCTCGAATCTTTCGGATGCGATGTTTCGATCGGTGACATTCCTCAA	600		
2511	GTGAGTGAGAACATCTCTCGAATCTTTCGGATGCGATGTTTCGATCGGTGACATTCCTCAA	2452		
601	CTCGTATCTCTGGATTTGCTCGCCACTTACCTAATTCGGGTGGTCTGATGTAGCAAG	660		
2451	CTCGTATCTCTGGATTTGCTCGCCACTTACCTAATTCGGGTGGTCTGATGTAGCAAG	2392		
661	ATCGCGGGTGAAGCTGATCAAGTTCTCGTGGGTATCTGACGAAATTTGCACTCTGGCGTT	720		
2391	ATCGCGGGTGAAGCTGATCAAGTTCTCGTGGGTATCTGACGAAATTTGCACTCTGGCGTT	2332		
721	GGTATGAGGTTTGGTATGTCGCGAAGGAAAGTGGTTCGATGCGAACCGTGTGACCTGCG	780		
2331	GGTATGAGGTTTGGTATGTCGCGAAGGAAAGTGGTTCGATGCGAACCGTGTGACCTGCG	2272		
781	GAAATCGCTGTTGATGAGTGGCGGGATCCACAGCGTTGCAATCTGTGGCGAGAAATTT	840		
2271	GAAATCGCTGTTGATGAGTGGCGGGATCCACAGCGTTGCAATCTGTGGCGAGAAATTT	2212		
841	GAGTTTCGTTTCGATGGCAAGTCGCGGCAATCGCGTGGTCCCGTGGATTCGGTGGCCAGCT	900		
2211	GAGTTTCGTTTCGATGGCAAGTCGCGGCAATCGCGTGGTCCCGTGGATTCGGTGGCCAGCT	2152		
901	GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGACGAGGAAGATCTGCCCGGTC	960		
2151	GGTCTTGGGGCAGAACTAACAGATGCTCAGATCGTTGACGAGGAAGATCTGCCCGGTC	2092		

```
Db 2571 ACTCAGGAAAAAGCGCTGGCAGCTCCAGCTTCCAGCTACTCAATGTTCAAGTGTGAC 2512
Qy 541 GTGAGTGAGAAATCACTCTCAATCTTCTCGATGCGATGTTGATCGGTGGACTTCCAAA 600
Db 2511 GTGAGTGAGAAATCACTCTCAATCTTCTCGATGCGATGTTGATCGGTGGACTTCCAAA 2452
Qy 601 CTCGTATCTCTGGATTTGCTGGCCACTACGTAAATTCGGGTGGTCTCGATGACGAAG 660
Db 2451 CTCGTATCTCTGGATTTGCTGGCCACTACGTAAATTCGGGTGGTCTCGATGACGAAG 2392
Qy 661 ATCGCGGTGAGCTGATCAAGTCTCTCGTGGTATCTGACGAAATTCATCTGCGCTT 720
Db 2391 ATCGCGGTGAGCTGATCAAGTCTCTCGTGGTATCTGACGAAATTCATCTGCGCTT 2332
Qy 721 GGTATGGAGTTGGTAGTGGCGACGGAAGAAAGTGGTGGATGGCAATGGCAACCTCTGG 780
Db 2331 GGTATGGAGTTGGTAGTGGCGACGGAAGAAAGTGGTGGATGGCAATGGCAACCTCTGG 2272
Qy 781 GAAATCGCTGTTGATGCACTGGCGGGGATCCACAAAGCGTTGGAACTGTGGCGAGAATTT 840
Db 2271 GAAATCGCTGTTGATGCACTGGCGGGGATCCACAAAGCGTTGGAACTGTGGCGAGAATTT 2212
Qy 841 GAGTTGGTTCGATGGGACGTTCGGCAATTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 900
Db 2211 GAGTTGGTTCGATGGGACGTTCGGCAATTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2152
Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATGCTGAGATGCTGAGCGAGAAATTCGCGG 960
Db 2151 GGTCTTGGGCGAGAACTAACAGATGCTCAGATGCTGAGCGAGAAATTCGCGGCGGTC 2092
Qy 961 ATGGTTGGGATCATTCGCGCGGATCGTGGATGATGATTCGAGCTTGTGGCGCTTACGTC 1020
Db 2091 ATGGTTGGGATCATTCGCGCGGATCGTGGATGATGATTCGAGCTTGTGGCGCTTACGTC 2032
Qy 1021 TTCGCGAGATCCTCGGATCTCGAAGCTCGGCGGATTCGGGAAATTCCTCGTATCAC 1080
Db 2031 TTCGCGAGATCCTCGGATCTCGAAGCTCGGCGGATTCGGGAAATTCCTCGTATCAC 1972
Qy 1081 TTGCAATTCGATTCGCGCGGATGCTGGCGCGGATGCTGGCGCGGATTCGCGCGGATTCG 1140
Db 1971 TTGCAATTCGATTCGCGCGGATGCTGGCGCGGATGCTGGCGCGGATTCGCGCGGATTCG 1912
```

RESULT 6

```
US-10-415-562A-5/c
; Sequence 5, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-415-562A-5
```

```
Query Match 100.0%; Score 1140; DB 17; Length 6334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACCGGTAAGTGTGAACACTTTCGGCGAAGACCGGCTCCCGTCTCTCGTGTGCG 60
Db 3051 ATGACCGGTAAGTGTGAACACTTTCGGCGAAGACCGGCTCCCGTCTCTCGTGTGCG 2992
Qy 61 TCCGATAGCGCGGATCCGGCGAGAACTGCGACCCAACTTCAACAAATCACCACGTC 120
```

```
Db 2991 TCCGATAGCGCGGATCCGGCGAGAACTGCGACCCAACTTCAACAAATCACCACGTC 2932
Qy 121 GAAACATTTAAACCGCTGTGGCGGGCGGATTTCTTGGCGTGAACCGGTGTGACCATTTGTCAAC 180
Db 2931 GAAACATTTAAACCGCTGTGGCGGGCGGATTTCTTGGCGTGAACCGGTGTGACCATTTGTCAAC 2872
Qy 181 GGTCCGAAGGTTCTGGATTCGGAGGCTTCTGTTCTCGGAAAGGGCTGGATCTGCCCC 240
Db 2871 GGTCCGAAGGTTCTGGATTCGGAGGCTTCTGTTCTCGGAAAGGGCTGGATCTGCCCC 2812
Qy 241 TGTCTGCGGGAAAGTGTGGTGCACATGTCGACGAAATTTCTCAAGTGTGTCAT 300
Db 2811 TGTCTGCGGGAAAGTGTGGTGCACATGTCGACGAAATTTCTCAAGTGTGTCAT 2752
Qy 301 CAACTCGGACTGGATCTGTTGCGATGTTGACGATGACCATCGCGCATACAGCTGGTCAG 360
Db 2751 CAACTCGGACTGGATCTGTTGCGATGTTGACGATGACCATCGCGCATACAGCTGGTCAG 2692
Qy 361 CGGCTCCAGGACTATGACCTGGATTTTCGGCAGCTTGGAAAGCTTGGCAACCAACGTCGT 420
Db 2691 CGGCTCCAGGACTATGACCTGGATTTTCGGCAGCTTGGAAAGCTTGGCAACCAACGTCGT 2632
Qy 421 CGTTCGGCTTACGGAACGTTGAAATGTACCGCTGCGACGATACGTCGGCGCTGTTGAATC 480
Db 2631 CGTTCGGCTTACGGAACGTTGAAATGTACCGCTGCGACGATACGTCGGCGCTGTTGAATC 2572
Qy 481 ACTCAGGAAAAAAGCGCTGGCAGCTCCAGCTTACGCGCTACTCATGTTCAAGTGGTGAC 540
Db 2571 ACTCAGGAAAAAAGCGCTGGCAGCTCCAGCTTACGCGCTACTCATGTTCAAGTGGTGAC 2512
Qy 541 GTGAGTGAGAACTCCTCGAATCTTTCGAGATGCGATGTTGATCGGTGGAATTCCTCAAA 600
Db 2511 GTGAGTGAGAACTCCTCGAATCTTTCGAGATGCGATGTTGATCGGTGGAATTCCTCAAA 2452
Qy 601 CTCGTATCTCTGGGATTTGCTGGCGCACTACGTAATTCGGGTGGTCTCGATGACGAAG 660
Db 2451 CTCGTATCTCTGGGATTTGCTGGCGCACTACGTAATTCGGGTGGTCTCGATGACGAAG 2392
Qy 661 ATCGCGGTGAGCTGATCAAGTTCCTCGCTCGTATCTGACGAAATTCGATCTGGCGTT 720
Db 2391 ATCGCGGTGAGCTGATCAAGTTCCTCGCTCGTATCTGACGAAATTCGATCTGGCGTT 2332
Qy 721 GGTATGAGGTTGTTAGTGGCGAGAAAGTGGTGGATGCGAACCCGTCACCTCTGG 780
Db 2331 GGTATGAGGTTGTTAGTGGCGAGAAAGTGGTGGATGCGAACCCGTCACCTCTGG 2272
Qy 781 GAAATCGCTGTTGATGCACTGGCGGGGATCCACAAAGCGTTGGAACTGTGGCGAGAATTT 840
Db 2271 GAAATCGCTGTTGATGCACTGGCGGGGATCCACAAAGCGTTGGAACTGTGGCGAGAATTT 2212
Qy 841 GAGTTTGTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGGATTCGCTGCCCGAGCT 900
Db 2211 GAGTTTGTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGGATTCGCTGCCCGAGCT 2152
Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATGCTGAGCGAGAAATTCGCGCGGTC 960
Db 2151 GGTCTTGGGCGAGAACTAACAGATGCTCAGATGCTGAGCGAGAAATTCGCGCGGTC 2092
Qy 961 ATGGTTGCGATCATTCGCGCGGATCGTGGATGATGATTCGAGCTTGTGGCGCTTACGTC 1020
Db 2091 ATGGTTGCGATCATTCGCGCGGATCGTGGATGATGATTCGAGCTTGTGGCGCTTACGTC 2032
Qy 1021 TTCGCGGAGATCCTCGGACTGTCGAAAGCTGGCGGACTTGGGAAATTCCTCGTATCAC 1080
Db 2031 TTCGCGGAGATCCTCGGACTGTCGAAAGCTGGCGGACTTGGGAAATTCCTCGTATCAC 1972
Qy 1081 TTGCAATTCGATTCGCGCGGATGCTGGCGCGGATGCTGGCGCGGATTCGCGCGGATTCG 1140
Db 1971 TTGCAATTCGATTCGCGCGGATGCTGGCGCGGATGCTGGCGCGGATTCGCGCGGATTCG 1912
```

RESULT 7

US-10-007-527A-7/c
; Sequence 7, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-527A-7

Query Match 100.0%; Score 1140; DB 14; Length 9652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGACACAGCGTAGTGCTGAAACACCTTTCGGCGAAAGACCGGCTCCCGTCTCGTGTCG 60
Db	5052	ATGACACAGCGTAGTGCTGAAACCTTTCGGCGAAAGACCGGCTCCCGTCTCGTGTCG 4993
Qy	61	TCCGATTAAGCGCGGCATCCGGCAGCAACTCGCACCCCAAACTTCAACAAATCACACGTCA 120
Db	4992	TCCGATTAAGCGCGGCATCCGGCAGCAACTCGCACCCCAAACTTCAACAAATCACACGTCA 4933
Qy	121	GAACATTTAACGCTGTGGCGGCGCATTTCTGGCGTGAACGGTGTGACCATGTGCAAC 180
Db	4932	GAACATTTAACGCTGTGGCGGCGCATTTCTGGCGTGAACGGTGTGACCATGTGCAAC 4873
Qy	181	GGTCCGAAAGTTCTCGAATTCGAGGCGCTTCGTTCTCGCGAAAGGGCTGGATCTGCCCC 240
Db	4872	GGTCCGAAAGTTCTCGAATTCGAGGCGCTTCGTTCTCGCGAAAGGGCTGGATCTGCCCC 4813
Qy	241	TGCTGTGCGGAAAGTCGCTGTCATCTGTGGCGTGAACGATGACATGCGCCATACAGCTGGTCAG 300
Db	4812	TGCTGTGCGGAAAGTCGCTGTCATCTGTGGCGTGAACGATGACATGCGCCATACAGCTGGTCAT 4753
Qy	301	CAACTCGGAGCTGGATCTGTTGGCATGCTGACGATGACCATGCGCCATACAGCTGGTCAG 360
Db	4752	CAACTCGGAGCTGGATCTGTTGGCATGCTGACGATGACCATGCGCCATACAGCTGGTCAG 4693
Qy	361	CGGCTCCAACGACCTATGAGCTGGATTTCCGACGCTTGGAAAGCTGCGACCAACGCTCGT 420
Db	4692	CGGCTCCAACGACCTATGAGCTGGATTTCCGACGCTTGGAAAGCTGCGACCAACGCTCGT 4633
Qy	421	CGTTGGCGGTACGGAACGTGAATCTGACGGCTGCGACGATACGTCGCGCTGTTGAAATC 480
Db	4632	CGTTGGCGGTACGGAACGTGAATCTGACGGCTGCGACGATACGTCGCGCTGTTGAAATC 4573
Qy	481	ACTCACGGAATAAAGCGGTGGCAGCTCCACGTTCCACGGCTACTCATGTTCACTGGTGCAG 540
Db	4572	ACTCACGGAATAAAGCGGTGGCAGCTCCACGTTCCACGGCTACTCATGTTCACTGGTGCAG 4513
Qy	541	GTGAGTGAGACATCTCGAATCCTTCTCGATCGGATGGATTCGATCGGTGACTTCCAAA 600
Db	4512	GTGAGTGAGACATCTCGAATCCTTCTCGATCGGATGGATTCGATCGGTGACTTCCAAA 4453
Qy	601	CTCGTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGGTGGTCTCGATGTACGAAG 660
Db	4452	CTCGTATCTCTGGGATTTGCTGCGCCACTACGTAATTCGGGTGGTCTCGATGTACGAAG 4393
Qy	661	ATCGCGGTGAGCTGATCAAGTTCTCGTGGTGTATCGACGAAAAATTGCATCTGGCGTT 720
Db	4392	ATCGCGGTGAGCTGATCAAGTTCTCGTGGTGTATCGACGAAAAATTGCATCTGGCGTT 4333

```
QY 241 TGCTGCGGAAAGTCGGTGCACATCGTGCAGCAAAATTTCTCAAGTTGTGCTCAT 300
Db |||||||
QY 4812 TGCTGTGCGGAAAGTCGGTGCACATCGTGCAGCAAAATTTCTCAAGTTGTGCTCAT 4753
Db |||||||
QY 301 CAATCGGGAGTGGATCTGTTGCCATGGTGCACGATGACCATGCCCATACAGCTGGTCAG 360
Db |||||||
QY 4752 CAATCGGGAGTGGATCTGTTGCCATGGTGCACGATGACCATGCCCATACAGCTGGTCAG 4693
Db |||||||
QY 361 CGGCTCCACGACCTATGAGCTGGATCTTTCGGCAGCCTGGAAAGTGGCAACACGGTCGT 420
Db |||||||
QY 4692 CGGCTCCACGACCTATGAGCTGGATCTTTCGGCAGCCTGGAAAGTGGCAACACGGTCGT 4633
Db |||||||
QY 421 CGTTGGCGTACGGAACTGAAATGATGAGCTGCGCTGCACGATACGCGCTGTTGAAATC 480
Db |||||||
QY 4632 CGTTGGCGTACGGAACTGAAATGATGAGCTGCGCTGCACGATACGCGCTGTTGAAATC 4573
Db |||||||
QY 481 ACTCACGAAACACGGCTGGCAGCTCCACGTTTCAACGCGTACTCATGTTCAAGTGGTGAC 540
Db |||||||
QY 4572 ACTCACGAAACACGGCTGGCAGCTCCACGTTTCAACGCGTACTCATGTTCAAGTGGTGAC 4513
Db |||||||
QY 541 GTGAGTGAGAACATCTCGAATCTTTCGGAATGCGATGTTTCGATCGGTGACATTCGAA 600
Db |||||||
QY 4512 GTGAGTGAGAACATCTCGAATCTTTCGGAATGCGATGTTTCGATCGGTGACATTCGAA 4453
Db |||||||
QY 601 CTGCTATCTCGGATTTGCTGCGGCACTACGTAATTCGGGTGCTCGATGATGACAAAG 660
Db |||||||
QY 4452 CTGCTATCTCGGATTTGCTGCGGCACTACGTAATTCGGGTGCTCGATGATGACAAAG 4393
Db |||||||
QY 661 ATCGGCGGTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTTGCAATCGCGT 720
Db |||||||
QY 4392 ATCGGCGGTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTTGCAATCGCGT 4333
Db |||||||
QY 721 GGTATGAGGTTGTTAGTGGCGACGGAACGAAAGTGGTGCACATCGGTCGACATTT 780
Db |||||||
QY 4332 GGTATGAGGTTGTTAGTGGCGACGGAACGAAAGTGGTGCACATCGGTCGACATTT 4273
Db |||||||
QY 781 GAATTCGCTGTTGATGCGAGTGGCGGATCCACAGCGTTTGAATCTGTGCGGAGAAATTT 840
Db |||||||
QY 4272 GAATTCGCTGTTGATGCGAGTGGCGGATCCACAGCGTTTGAATCTGTGCGGAGAAATTT 4213
Db |||||||
QY 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTCGTCGCCGAGCT 900
Db |||||||
QY 4212 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTCGTCGCCGAGCT 4153
Db |||||||
QY 901 GGTCTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAATTCGCCCGGTC 960
Db |||||||
QY 4152 GGTCTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAATTCGCCCGGTC 4093
Db |||||||
QY 961 ATGTTGCGATCATTCGGGCGGATCGTGGATGATGATTCGACCTTGTGCGCCTTACGTC 1020
Db |||||||
QY 4092 ATGTTGCGATCATTCGGGCGGATCGTGGATGATGATTCGACCTTGTGCGCCTTACGTC 4033
Db |||||||
QY 1021 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTCGATCAC 1080
Db |||||||
QY 4032 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTCGATCAC 3973
Db |||||||
QY 1081 TTGCATTATCATTTGCCGAGCGATGTCGGGCCCCCGATATATTCGGTTTCGCAAGTGA 1140
Db |||||||
QY 3972 TTGCATTATCATTTGCCGAGCGATGTCGGGCCCCCGATATATTCGGTTTCGCAAGTGA 3913
Db |||||||
```

RESULT 9

```
US-10-415-562A-7/c
; Sequence 7, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CLI709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
```

```
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHRL7
US-10-415-562A-7
```

```
Query Match 100.0%; Score 1140; DB 17; Length 9652;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGACACAGCTGAAGTGTGAACACCTTTCCGGGAAAGACCGGCTCCCGCTCTCGTGTGCG 60
Db |||||||
QY 5052 ATGACACAGCTGAAGTGTGAACACCTTTCCGGGAAAGACCGGCTCCCGCTCTCGTGTGCG 4993
Db |||||||
QY 61 TCCGATAAGCGCGGCATCCGCGACGAACTGCGACCCAACTTCAACAATACACACGTCA 120
Db |||||||
QY 4992 TCCGATAAGCGCGGCATCCGCGACGAACTGCGACCCAACTTCAACAATACACACGTCA 4933
Db |||||||
QY 121 GAAACATTTAACGCTCTGTGGCGGCGGATTTCTGCGGTGACGATGACATTTGTCAAC 180
Db |||||||
QY 4932 GAAACATTTAACGCTCTGTGGCGGCGGATTTCTGCGGTGACGATGACATTTGTCAAC 4873
Db |||||||
QY 181 GGTCCGAAAGCTTCTGATTTCGGAGGCTTCTGTTCCGCGAAAGGCTGATCTGCCCC 240
Db |||||||
QY 4872 GGTCCGAAAGCTTCTGATTTCGGAGGCTTCTGTTCCGCGAAAGGCTGATCTGCCCC 4813
Db |||||||
QY 241 TGCTGTGCGGAAAGTTCGGTGCAATCGTGCAGCGAAATTTCTCAAGTTGTTGCTCAT 300
Db |||||||
QY 4812 TGCTGTGCGGAAAGTTCGGTGCAATCGTGCAGCGAAATTTCTCAAGTTGTTGCTCAT 4753
Db |||||||
QY 301 CAATCGGGACTGATCTGTTGCGATGCGTGCAGTACCATGCGCCATACAGCTGGTCAG 360
Db |||||||
QY 4752 CAATCGGGACTGATCTGTTGCGATGCGTGCAGTACCATGCGCCATACAGCTGGTCAG 4693
Db |||||||
QY 361 CGGCTCCACGACCTATGAGCTGGATCTTTCGGCAGCCTGGAAAGTGGCAACACGGTCGT 420
Db |||||||
QY 4692 CGGCTCCACGACCTATGAGCTGGATCTTTCGGCAGCCTGGAAAGTGGCAACACGGTCGT 4633
Db |||||||
QY 421 CGTTGGCGTACGGAACTGAAATGATGAGCTGCGCTGCACGATGATGACAAATCTTGAATC 480
Db |||||||
QY 4632 CGTTGGCGTACGGAACTGAAATGATGAGCTGCGCTGCACGATGATGACAAATCTTGAATC 4573
Db |||||||
QY 481 ACTCACGAAACACGGCTGGCAGCTCCACGTTTCAACGCGTACTCATGTTCAAGTGGTGAC 540
Db |||||||
QY 4572 ACTCACGAAACACGGCTGGCAGCTCCACGTTTCAACGCGTACTCATGTTCAAGTGGTGAC 4513
Db |||||||
QY 541 GTGAGTGAGAACATCTCGAATCTTTCGGAATGCGATGTTTCGATCGGTGACATTCGAA 600
Db |||||||
QY 4512 GTGAGTGAGAACATCTCGAATCTTTCGGAATGCGATGTTTCGATCGGTGACATTCGAA 4453
Db |||||||
QY 601 CTCGCTATCTCGGATTTGCTGCGGCACTACGTAATTCGGGTGCTCGATGATGACAAAG 660
Db |||||||
QY 4452 CTCGCTATCTCGGATTTGCTGCGGCACTACGTAATTCGGGTGCTCGATGATGACAAAG 4393
Db |||||||
QY 661 ATCGGCGGTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTTGCAATCGCGT 720
Db |||||||
QY 4392 ATCGGCGGTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTTGCAATCGCGT 4333
Db |||||||
QY 721 GGTATGAGGTTGTTAGTGGCGACGGAACGAAAGTGGTGCACATCGGTCGACATTT 780
Db |||||||
QY 4332 GGTATGAGGTTGTTAGTGGCGACGGAACGAAAGTGGTGCACATCGGTCGACATTT 4273
Db |||||||
QY 781 GAAATTCGCTGTTGATGCGAGTGGCGGATCCACAGCGTTTGAATCTGTGCGGAGAAATTT 840
Db |||||||
QY 4272 GAAATTCGCTGTTGATGCGAGTGGCGGATCCACAGCGTTTGAATCTGTGCGGAGAAATTT 4213
Db |||||||
QY 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTCGTCGCCGAGCT 900
Db |||||||
QY 4212 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCTGGTCCCGTGGATTCGTCGCCGAGCT 4153
Db |||||||
QY 901 GGTCTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAATTCGCCCGGTC 960
Db |||||||
QY 4152 GGTCTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAATTCGCCCGGTC 4093
Db |||||||
QY 961 ATGTTGCGATCATTCGGGCGGATCGTGGATGATGATTCGACCTTGTGCGCCTTACGTC 1020
Db |||||||
QY 4092 ATGTTGCGATCATTCGGGCGGATCGTGGATGATGATTCGACCTTGTGCGCCTTACGTC 4033
Db |||||||
QY 1021 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTCGATCAC 1080
Db |||||||
QY 4032 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTCGATCAC 3973
Db |||||||
QY 1081 TTGCATTATCATTTGCCGAGCGATGTCGGGCCCCCGATATATTCGGTTTCGCAAGTGA 1140
Db |||||||
QY 3972 TTGCATTATCATTTGCCGAGCGATGTCGGGCCCCCGATATATTCGGTTTCGCAAGTGA 3913
Db |||||||
```

```
Db 4152 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGACGAGAAATCTCCCGCGTC 4093
QY 961 ATGGTTGGGATCATTTCCGGCGGAGATCGTGGATGATGATTCGGACTTGGCGCTTACGTC 1020
Db 4092 ATGGTTGGGATCATTTCCGGCGGAGATCGTGGATGATGATTCGGACTTGGCGCTTACGTC 4033
QY 1021 TTCGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 1080
Db 4032 TTCGCGAGATCCTCGGACTCGTCGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 3973
QY 1081 TTGCATTATCGATTGCCCGCAGCGGATGTCGCGCCCCCGGATTAATTCGGTTCGCAAGTGA 1140
Db 3972 TTGCATTATCGATTGCCCGCAGCGGATGTCGCGCCCCCGGATTAATTCGGTTCGCAAGTGA 3913

RESULT 10
US-10-007-527A-6/c
; Sequence 6, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CLI709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid prHBR17
US-10-007-527A-6

Query Match 100.0%; Score 1140; DB 14; Length 11241;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAGTAAAGTGTGAACACCTTTCGGGCAAGACCGGCTCCCGTCTCGTGTGCG 60
Db 6641 ATGACGAGTAAAGTGTGAACACCTTTCGGGCAAGACCGGCTCCCGTCTCGTGTGCG 6582

QY 61 TCCGATAAGCGCGGCATCCGGACGAACTGGGACCCAAACTTCAACAAATCACCACGTCA 120
Db 6581 TCCGATAAGCGCGGCATCCGGACGAACTGGGACCCAAACTTCAACAAATCACCACGTCA 6522

QY 121 GAACATTTAAGCTGTGGCGGCGGATTTCTGCGGTGAACGGTGTGACCATTTGTCAAC 180
Db 6521 GAACATTTAAGCTGTGGCGGCGGATTTCTGCGGTGAACGGTGTGACCATTTGTCAAC 6462

QY 181 GGTCCGAAAGTTCGTGAATTCGGAGCGCTTCGTTCCTCGGAAAGGGCTGGATCTGCCCC 240
Db 6461 GGTCCGAAAGTTCGTGAATTCGGAGCGCTTCGTTCCTCGGAAAGGGCTGGATCTGCCCC 6402

QY 241 TGCTGTGGGAAAGTTCGTGCAATCGTGACAGCAAAATTTCTCAAGTTGTGTCAT 300
Db 6401 TGCTGTGGGAAAGTTCGTGCAATCGTGACAGCAAAATTTCTCAAGTTGTGTCAT 6342

QY 301 CAACTCGGAGCTGATCTGTTGCGATGTTGACATGACCATGCGCATACAGCTGTGAC 360
Db 6341 CAACTCGGAGCTGATCTGTTGCGATGTTGACATGACCATGCGCATACAGCTGTGAC 6282

QY 361 CGGCTCCACGACCTATGACTTGGACTTTTCGGCAGCTCGAAAGCTGCGACCAACGGTGGT 420
Db 6281 CGGCTCCACGACCTATGACTTGGACTTTTCGGCAGCTCGAAAGCTGCGACCAACGGTGGT 6222

QY 421 CGTTGGGCTAGGAACTGAAATGATCGGCTCGACGATACGTGCGCGCTGTGGAATC 480
```

```
Db 6221 CGTTGGGCTAGGAACTGAAATGATACGGCTCGACGGATACGTGGCGCTGTGGAATC 6162
QY 481 ACTACGGAATAAAACGGCTGGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 540
Db 6161 ACTACGGAATAAAACGGCTGGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 6102
QY 541 GTGAGTGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTTGGAATTCCTCAA 600
Db 6101 GTGAGTGAACATCCTCGAATCCTTCTCGGATCGATGTTTCGATCGGTTGGAATTCCTCAA 6042
QY 601 CTCGTATCTCTGGGATTTGCTGCGCACTACGTAATTCGGGTGGTCTCGATACGAAAG 660
Db 6041 CTCGTATCTCTGGGATTTGCTGCGCACTACGTAATTCGGGTGGTCTCGATACGAAAG 5982
QY 661 ATCGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGAACGAAATTCGATCTGCGGTT 720
Db 5981 ATCGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGAACGAAATTCGATCTGCGGTT 5922
QY 721 GGTATGGAGTTGGTGGCGAGGAAAGTGTGACATGCGCAACCGTGCACCCCTGG 780
Db 5921 GGTATGGAGTTGGTGGCGAGGAAAGTGTGACATGCGCAACCGTGCACCCCTGG 5862
QY 781 GAAATCGCTGTGATGCAAGTGGCGGGGATCCAAAGCGTTTGGAACTGTGGCGAGAAATTT 840
Db 5861 GAAATCGCTGTGATGCAAGTGGCGGGGATCCAAAGCGTTTGGAACTGTGGCGAGAAATTT 5802
QY 841 GAGTTGGTTCGATGGGACGTGCGGCAATCGGCTGCTGCGGTGGATTCGCTGCCGAGCT 900
Db 5801 GAGTTGGTTCGATGGGACGTGCGGCAATCGGCTGCTGCGGTGGATTCGCTGCCGAGCT 5742
QY 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 960
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCGGTC 5682
QY 961 ATGTTGCGATCATTTCCGCGCGATCGTGGATGATGATTCGACCTTGTGCGCTTACGTC 1020
Db 5681 ATGTTGCGATCATTTCCGCGCGATCGTGGATGATGATTCGACCTTGTGCGCTTACGTC 5622
QY 1021 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGAGCTTGGGAAATCTTCGTGATCAC 1080
Db 5621 TTCGCGAGATCCTCGGACTCGTGAAGCTGGCGGAGCTTGGGAAATCTTCGTGATCAC 5562
QY 1081 TTGCATTATCGATTGCCCGCAGCGATGTCGCGCCCCCGGATTAATTCGCTTCGCAAGTGA 1140
Db 5561 TTGCATTATCGATTGCCCGCAGCGATGTCGCGCCCCCGGATTAATTCGCTTCGCAAGTGA 5502

RESULT 11
US-10-007-452-6/c
; Sequence 6, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CLI709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid prHBR17
US-10-007-452-6

Query Match 100.0%; Score 1140; DB 14; Length 11241;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGACACGCGTAAAGTGTGTAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGCTGCG 60
Db 6641 ATGACACGCGTAAAGTGTGTAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGCTGCG 6582

Qy 61 TCCGATAAGCGCGCATCCGCGACGAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 120
Db 6581 TCCGATAAGCGCGCATCCGCGACGAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 6522

Qy 121 GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 180
Db 6521 GAAACATTTAAACGCTGTGGCGCGCGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 6462

Qy 181 GGTCCGAAAGGTTCTGATTCGGAGGCTTCTGTTCTCGGAAAGGCTGATCTGCCCC 240
Db 6461 GGTCCGAAAGGTTCTGATTCGGAGGCTTCTGTTCTCGGAAAGGCTGATCTGCCCC 6402

Qy 241 TGCTGTCCGGGAAAGTCCGTCACATCGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db 6401 TGCTGTCCGGGAAAGTCCGTCACATCGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 6342

Qy 301 CAACTCGGGAATGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 360
Db 6341 CAACTCGGGAATGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 6282

Qy 361 CGGCTCCACAGCATGATGGAATTCGCGACGCTTTCGGAGGCTGGAAGCTGACCAACGGTGT 420
Db 6281 CGGCTCCACAGCATGATGGAATTCGCGACGCTTTCGGAGGCTGGAAGCTGACCAACGGTGT 6222

Qy 421 CGTTGGGCTACGGAACGTTGAGTGTGACGATGATGATGATGATGATGATGATGATGATG 480
Db 6221 CGTTGGGCTACGGAACGTTGAGTGTGACGATGATGATGATGATGATGATGATGATGATG 6162

Qy 481 ACTCAGCGAAACGCTGCGACGCTCACGTTTCAACGCTGATCATGTTTCAAGTGGTCA 540
Db 6161 ACTCAGCGAAACGCTGCGACGCTCACGTTTCAACGCTGATCATGTTTCAAGTGGTCA 6102

Qy 541 GTGAGTGAGAACATCTCTGGAATTCGCGTGTGACGATGATGATGATGATGATGATGATG 600
Db 6101 GTGAGTGAGAACATCTCTGGAATTCGCGTGTGACGATGATGATGATGATGATGATGATG 6042

Qy 601 CTGCTATCTCTGGGATTTGCTGCGGCTACGTAATTCGGGTGTCTCGATGATGATGATG 660
Db 6041 CTGCTATCTCTGGGATTTGCTGCGGCTACGTAATTCGGGTGTCTCGATGATGATGATGATG 5982

Qy 661 ATCGCGGCTGAAGCTGATCAAGTTCTCGTCTGCTATCTGACGAAATTTGCAATCTGGGTT 720
Db 5981 ATCGCGGCTGAAGCTGATCAAGTTCTCGTCTGCTATCTGACGAAATTTGCAATCTGGGTT 5922

Qy 721 GGTATGAGGTTGATGAGGACGGAAGAAAGTGGTGAACATGCGACATGCGACCCCTGG 780
Db 5921 GGTATGAGGTTGATGAGGACGGAAGAAAGTGGTGAACATGCGACATGCGACCCCTGG 5862

Qy 781 GAAATCGCTGTGATGAGTGGGCGGATCCAAAGCGTTGGAACATGTCGCGGAGAAATTT 840
Db 5861 GAAATCGCTGTGATGAGTGGGCGGATCCAAAGCGTTGGAACATGTCGCGGAGAAATTT 5802

Qy 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCCGCTGGTCCGCTGATGATGATGATGATGATG 900
Db 5801 GAGTTTGGTTCGATGGGACGTCGGGCAATCCGCTGGTCCGCTGATGATGATGATGATGATG 5742

Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCCGGTC 960
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCCGGTC 5682

Qy 961 ATGTTTGGGATCATTTCCGCGCGATTCGTTGATGATGATGATGATGATGATGATGATGATG 1020
Db 5681 ATGTTTGGGATCATTTCCGCGCGATTCGTTGATGATGATGATGATGATGATGATGATGATG 5622

Qy 1021 TTCGGCGAGATCTCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTTGATCAC 1080
Db 5621 TTCGGCGAGATCTCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTTGATCAC 5562
```

```
Qy 1081 TTGCATTATCGATTGCCCGCAGCGGATGTGGGCCCCCGGATAATATATCGTTCGCAAGTGA 1140
Db 5561 TTGCATTATCGATTGCCCGCAGCGGATGTGGGCCCCCGGATAATATCGTTCGCAAGTGA 5502

RESULT 12
US-10-415-562A-6/c
; Sequence 6, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid pRHB17
US-10-415-562A-6
```

```
Query Match 100.0%; Score 1140; DB 17; Length 11241;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACACGCGTAAAGTGTGTAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGCTGCG 60
Db 6641 ATGACACGCGTAAAGTGTGTAACACCTTTCGGGAAAGACCGGCTCCGGTCTCTGCTGCG 6582

Qy 61 TCCGATAAGCGCGCATCCGCGACGAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 120
Db 6581 TCCGATAAGCGCGCATCCGCGACGAACTCGGACCCCAAACTTCAACAAATCACACAGTCA 6522

Qy 121 GAAACATTTAAACGCTGTGCGCGCGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 180
Db 6521 GAAACATTTAAACGCTGTGCGCGCGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 6462

Qy 181 GGTCCGAAAGGTTCTGATTCGGAGGCTTCTGTTCTCGGAAAGGCTGATCTGCCCC 240
Db 6461 GGTCCGAAAGGTTCTGATTCGGAGGCTTCTGTTCTCGGAAAGGCTGATCTGCCCC 6402

Qy 241 TGCTGTCCGGGAAAGTCCGTCACATCGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 300
Db 6401 TGCTGTCCGGGAAAGTCCGTCACATCGTCAGACGAAATTTCTCAAGTTGTTGCTCAT 6342

Qy 301 CAACTCGGGAATGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 360
Db 6341 CAACTCGGGAATGATCTGTTGCGATGGTGAACGATGACCAATGCGCCATACAGCTGGTCA 6282

Qy 361 CGGCTCCACAGCATGATGGAATTCGCGACGCTTTCGGAGGCTGGAAGCTGACCAACGGTGT 420
Db 6281 CGGCTCCACAGCATGATGGAATTCGCGACGCTTTCGGAGGCTGGAAGCTGACCAACGGTGT 6222

Qy 421 GGTGGCGTACGGAACGTTGAGTGTGACGATGATGATGATGATGATGATGATGATGATG 480
Db 6221 GGTGGCGTACGGAACGTTGAGTGTGACGATGATGATGATGATGATGATGATGATGATG 6162

Qy 481 ACTCAGCGAAACGCTGCGACGCTCACGTTTCAACGCTGATCATGTTTCAAGTGGTCA 540
Db 6161 ACTCAGCGAAACGCTGCGACGCTCACGTTTCAACGCTGATCATGTTTCAAGTGGTCA 6102

Qy 541 GTGAGTGAGAACATCTCTGGAATTCGCGTGTGACGATGATGATGATGATGATGATGATG 600
Db 6101 GTGAGTGAGAACATCTCTGGAATTCGCGTGTGACGATGATGATGATGATGATGATGATG 6042

Qy 601 CTGCTATCTCTGGGATTTGCTGCGGCTACGTAATTCGGGTGTCTCGATGATGATGATG 660
Db 6041 CTGCTATCTCTGGGATTTGCTGCGGCTACGTAATTCGGGTGTCTCGATGATGATGATGATG 5982

Qy 661 ATCGCGGCTGAAGCTGATCAAGTTCTCGTCTGCTATCTGACGAAATTTGCAATCTGGGTT 720
Db 5981 ATCGCGGCTGAAGCTGATCAAGTTCTCGTCTGCTATCTGACGAAATTTGCAATCTGGGTT 5922

Qy 721 GGTATGAGGTTGATGAGGACGGAAGAAAGTGGTGAACATGCGACATGCGACCCCTGG 780
Db 5921 GGTATGAGGTTGATGAGGACGGAAGAAAGTGGTGAACATGCGACATGCGACCCCTGG 5862

Qy 781 GAAATCGCTGTGATGAGTGGGCGGATCCAAAGCGTTGGAACATGTCGCGGAGAAATTT 840
Db 5861 GAAATCGCTGTGATGAGTGGGCGGATCCAAAGCGTTGGAACATGTCGCGGAGAAATTT 5802

Qy 841 GAGTTTGGTTCGATGGGACGTCGGGCAATCCGCTGGTCCGCTGATGATGATGATGATGATG 900
Db 5801 GAGTTTGGTTCGATGGGACGTCGGGCAATCCGCTGGTCCGCTGATGATGATGATGATGATG 5742

Qy 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCCGGTC 960
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAGAAATCTGCCCCGGTC 5682

Qy 961 ATGTTTGGGATCATTTCCGCGCGATTCGTTGATGATGATGATGATGATGATGATGATGATG 1020
Db 5681 ATGTTTGGGATCATTTCCGCGCGATTCGTTGATGATGATGATGATGATGATGATGATGATG 5622

Qy 1021 TTCGGCGAGATCTCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTTGATCAC 1080
Db 5621 TTCGGCGAGATCTCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTTCGTTGATCAC 5562
```

QY 661 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCATCTCGCGTT 720
Db |||||
QY 5981 ATCGCGGTGAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCATCTCGCGTT 5922
Db |||||
QY 721 GGTATGGAGTTGGTAGTGGCGACGAAAGTGGTGCACATGGCAACCGTGCACCCCTGG 780
Db |||||
QY 5921 GGTATGGAGTTGGTAGTGGCGACGAAAGTGGTGCACATGGCAACCGTGCACCCCTGG 5862
Db |||||
QY 781 GAAATCGCTGTTGATGCAATGCGGCGGATCCACAAAGCGTTGGAATCTGCGGAGAAATT 840
Db |||||
QY 5861 GAAATCGCTGTTGATGCAATGCGGCGGATCCACAAAGCGTTGGAATCTGCGGAGAAATT 5802
Db |||||
QY 841 GAGTTTGGTTCATCGGAGCTCGGCAATCGCGTGGTCCCGTGGATTCGGTCCCGAGCT 900
Db |||||
QY 5801 GAGTTTGGTTCATCGGAGCTCGGCAATCGCGTGGTCCCGTGGATTCGGTCCCGAGCT 5742
Db |||||
QY 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGTCTCCCGCGTTC 960
Db |||||
QY 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAAGTCTCCCGCGTTC 5682
Db |||||
QY 961 ATGGTTGGCATTCATCCGCGCGATCGTGGATGATGATTCGGAATTCGCGCTTACGTC 1020
Db |||||
QY 5681 ATGGTTGGCATTCATCCGCGCGATCGTGGATGATGATTCGGAATTCGCGCTTACGTC 5622
Db |||||
QY 1021 TTCGCGAGATCTCGGACTCTCGAAGCTGCGGAGCTGCGGAAATTCGTCATCAC 1080
Db |||||
QY 5621 TTCGCGAGATCTCGGACTCTCGAAGCTGCGGAGCTGCGGAAATTCGTCATCAC 5562
Db |||||
QY 1081 TTGCATTCATTCGATTCGCGCGAGGATGCGCGCGCGCGATTAATTCGTCGCAAGTGA 1140
Db |||||
QY 5561 TTGCATTCATTCGATTCGCGCGAGGATGCGCGCGCGCGATTAATTCGTCGCAAGTGA 5502
Db |||||

RESULT 13
US-10-184-644-332
; Sequence 332, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 332
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-332

Query Match 3.6%; Score 40.8; DB 14; Length 520;
Best Local Similarity 18.5%; Pred. No. 0.018;
Matches 74; Conservative 81; Mismatches 246; Indels 0; Gaps 0;
QY 522 ACTCATGTTTCAGTGTGAGTGCAGTGCAGAACTCTCGAATCTCTCGATCGCATGTT 581
Db |||||
QY 90 DSPSFLQSAHPGHEHLAQASRLVQLQALTWVRVSHLQLRVNDFTQNPGRFRIKGE 149
Db |||||
QY 582 CGATCGGTGACTTCCAACTCGTATCTCGGATTTCTCGCCACTACGTAATTCGGG 641
Db |||||
QY 150 QGAPGLQKHGKMGMPGAPGPPPAEKAGKMGMRDGTGSPGPPGVKGEAGLQGP 209
Db |||||

QY 642 TGGTCTCGATGTACGAAAGATCGCGGTGAACTGATCAAGTTCTCGTGGTATCTGAC 701
Db |||||
QY 210 QGAPGKQATGTPGPQGEKSGKDGGLIGPKGETKGEKGLGLPGSKGDRGMKGDAGV 269
Db |||||
QY 702 GAAATTCATCTCGCGTTGGTATCGAGTTGGTAGTGGCGACGCAAAAGTGGTCGACA 761
Db |||||
QY 270 MGPPGAQSGKDFGRPPGPPGLAGFPFGAKDQGPGLQGVPPGAVHPGAKGEPGSGS 329
Db |||||
QY 762 TGGCAACCGTGCACCTCGGAAATCGCTGTTGATCAGTGGCGGGGATCCACAAAGCGTT 821
Db |||||
QY 330 PBRAGLPSPGSPGATGLKSGKDTGLGQQQGRKESGVPAGVKGQSGPLAGPKGA 389
Db |||||
QY 822 GGAATCTGGCGAGAAATTTGATTTGGTTCGATGGAGCGTCGGGCAATCGCTGTCGCG 881
Db |||||
QY 390 PQAGKQKDGQVKGSSGQGVKGEKRGENSEVSRIVGSSNRGAENVYVYSGTWGTCDD 449
Db |||||
QY 882 TGGATTGGTGGCGCGAGCTGCTTGGGCGCAGAACTAACAG 922
Db |||||
QY 450 EWQNSDALIVFCRMLGYSKRALYKVGAGTGQIWLNDNVQCRG 490
Db |||||

RESULT 14
US-10-184-634-332
; Sequence 332, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 332
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-332

Query Match 3.6%; Score 40.8; DB 14; Length 520;
Best Local Similarity 18.5%; Pred. No. 0.018;
Matches 74; Conservative 81; Mismatches 246; Indels 0; Gaps 0;
QY 522 ACTCATGTTTCAGTGTGAGTGCAGTGCAGAACTCTCGAATCTCTCGATCGCATGTT 581
Db |||||
QY 90 DSPSFLQSAHPGHEHLAQASRLVQLQALTWVRVSHLQLRVNDFTQNPGRFRIKGE 149
Db |||||
QY 582 CGATCGGTGACTTCCAACTCGTATCTCGGATTTCTCGCCACTACGTAATTCGGG 641
Db |||||
QY 150 QGAPGLQKHGKMGMPGAPGPPPAEKAGKMGMRDGTGSPGPPGVKGEAGLQGP 209
Db |||||
QY 642 TGGTCTCGATGTACGAAAGATCGCGGTGAACTGATCAAGTTCTCGTGGTATCTGAC 701
Db |||||
QY 210 QGAPGKQATGTPGPQGEKSGKDGGLIGPKGETKGEKGLGLPGSKGDRGMKGDAGV 269
Db |||||
QY 702 GAAATTCATCTCGCGTTGGTATCGAGTTGGTAGTGGCGACGCAAAAGTGGTCGACA 761
Db |||||
QY 270 MGPPGAQSGKDFGRPPGPPGLAGFPFGAKDQGPGLQGVPPGAVHPGAKGEPGSGS 329
Db |||||
QY 762 TGGCAACCGTGCACCTCGGAAATCGCTGTTGATCAGTGGCGGGGATCCACAAAGCGTT 821
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 02:18:21 ; Search time 85 Seconds
(without alignments)
1599.511 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005
Sequence: 1 MTSVSAHLSGKDRPPVLVS.....HLHYRLPAADVRPIISVRK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2005	100.0	379	5	ABB84278
2	609.5	30.4	459	5	ABB84280
3	399.5	19.9	494	2	AAW73071
4	326	16.3	427	3	AAW7536
5	326	16.3	427	5	ABB06341
6	326	16.3	427	5	AAE13530
7	322	16.1	451	5	ABB84283
8	318	15.9	427	3	AAW7537
9	318	15.9	427	5	ABB06342
10	318	15.9	427	5	AAE13531
11	318	15.9	427	7	ADB66179
12	316	15.8	456	5	ABB84281
13	282	14.1	466	5	AAU11039
14	266.5	13.3	528	5	ABB84282
15	134	6.7	314	1	AAW7537
16	104	5.2	409	6	AAE37298
17	103	5.1	647	6	ABU56572
18	101.5	5.1	882	7	ABO68907
19	101	5.0	1395	7	ABO81599
20	100.5	5.0	577	2	AAW06554
21	100	5.0	1568	4	ABG03109
22	99.5	5.0	676	7	ABO74435
23	98.5	4.9	532	6	ABU15733
24	98.5	4.9	621	7	ABO70485
25	98.5	4.9	1409	7	ABO73536

ALIGNMENTS

RESULT 1
ABB84278

ID ABB84278 standard; protein; 379 AA.

XX ABB84278;

XX AC ABB84278;

XX DT 13-JAN-2003 (first entry)

XX DE Rhodococcus AN12 replication protein Rep.

XX KW Plasmid stability protein; replication protein; ethylene forming enzyme;

XX KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;

XX KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;

XX KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;

XX KW shuttle vector; Rep protein.

XX OS Rhodococcus erythropolis.

XX EN WO200255709-A2.

XX PD 18-JUL-2002.

XX PF 12-DEC-2001; 2001WO-US047868.

XX PR 12-DEC-2000; 2000US-0254868P.

XX FA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;

XX WPI; 2002-557827/59.

XX DR N-PSDB; ABQ76122.

XX PT New nucleic acid molecule encoding replication protein/plasmid stability

XX PT protein, useful in cloning and expression vectors, particularly shuttle

XX PT vectors for expression of heterologous genes in Rhodococcus species.

XX PS Claim 4; Page 64-65; 96pp; English.

XX CC This invention describes a novel nucleic acid encoding a replication

XX CC protein or a plasmid stability protein. The product of the invention is

XX CC useful for expression of nucleic acid such as genes encoding enzymes

XX CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic

XX CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile

XX CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol

XX CC dehydrogenase, terpene synthases, and cholesterol oxidase in an

XX CC Actinomycetales bacteria. The replication protein or plasmid stability

XX CC protein are useful in cloning and expression vectors and particularly in

CC shuttle vectors for the expression of homologous and heterologous genes
 CC in *Rhodococcus* sp. This sequence represents the *Rhodococcus* AN12 Rep
 CC protein described in the disclosure of the invention
 XX
 XX
 SQ Sequence 379 AA;
 Query Match 100.0%; Score 2005; DB 5; Length 379;
 Best Local Similarity 100.0%; Pred. No. 6.6e-193; Mismatches 0; Indels 0; Gaps 0;
 Matches 379; Conservative 0;
 QY 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOQITTSFNAAGRPISGVNGVTIVN 60
 DB 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOQITTSFNAAGRPISGVNGVTIVN 60
 QY 61 GPKSGGFGGLRSCGKWNCPCCAGKVGHAHRADEISQVVAHQIGTGSVAMVTMTMHTAGQ 120
 DB 61 GPKSGGFGGLRSCGKWNCPCCAGKVGHAHRADEISQVVAHQIGTGSVAMVTMTMHTAGQ 120
 QY 121 RLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGWHVHVHALLMFSGD 180
 DB 121 RLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGWHVHVHALLMFSGD 180
 QY 181 VSENILESFSDAMFDRWTSKLVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYLTAKIASGV 240
 DB 181 VSENILESFSDAMFDRWTSKLVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYLTAKIASGV 240
 QY 241 GMEVSGGKSGRHNRAPEIAVDVGDPOALELWREPEFGSMGRRRAIWSRGLRARA 300
 DB 241 GMEVSGGKSGRHNRAPEIAVDVGDPOALELWREPEFGSMGRRRAIWSRGLRARA 300
 QY 301 GLGAELTDQAIVQEESAPVMVAIIPARSWMWIRTCAPYVFEIILGLVEAGATWENLRDH 360
 DB 301 GLGAELTDQAIVQEESAPVMVAIIPARSWMWIRTCAPYVFEIILGLVEAGATWENLRDH 360
 QY 361 LHYRLPAADVRRPPIISVRK 379
 DB 361 LHYRLPAADVRRPPIISVRK 379
 RESULT 2
 ID ABB84280 standard; protein; 459 AA.
 AC ABB84280;
 XX
 XX
 DT 13-JAN-2003 (first entry)
 DE
 DE A. pyogenes pAP1 replication protein.
 KW Plasmid stability protein; replication protein; ethylene forming enzyme;
 KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
 KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
 KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
 KW shuttle vector.
 XX
 OS Arcanobacterium pyogenes.
 XX
 XX WO200255709-A2.
 XX
 XX 18-JUL-2002.
 XX
 XX 12-DEC-2001; 2001WO-US047868.
 XX
 XX 12-DEC-2000; 2000US-0254868P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 XX Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
 XX
 XX WPI; 2002-557827/59.
 XX
 XX New nucleic acid molecule encoding replication protein/plasmid stability
 XX protein, useful in cloning and expression vectors, particularly shuttle

PT vectors for expression of heterologous genes in *Rhodococcus* species.
 XX
 PS Example 5; Fig 4A; 96pp; English.
 XX
 CC This invention describes a novel nucleic acid encoding a replication
 CC protein or a plasmid stability protein. The product of the invention is
 CC useful for expression of nucleic acid such as genes encoding enzymes
 CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
 CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
 CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
 CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
 CC Actinomycetales bacteria. The replication protein or plasmid stability
 CC protein are useful in cloning and expression vectors and particularly in
 CC shuttle vectors for the expression of homologous and heterologous genes
 CC in *Rhodococcus* sp. This sequence represents A. pyogenes pAP1 replication
 CC protein described in the disclosure of the invention
 XX
 SQ Sequence 459 AA;
 Query Match 30.4%; Score 609.5; DB 5; Length 459;
 Best Local Similarity 36.3%; Pred. No. 3.5e-52;
 Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;
 QY 5 SAEHLGKDRPPVLVSSD-----KRGIRHELPRKLOQITTSFNAAGRPISGVNGV 56
 DB 67 SATPLGNVTLTFPVSNESKKTAKRSRERYELDRGLAEITLESVRKGRVPVAPLVSL 126
 QY 57 TIVNPKSGGFGGLRSCGKWNCPCCAGKVGHAHRADEISQVVAHQIGTG-SVAMVTMTMR 115
 DB 127 RAKSDGKAGYGLHTCGSVWACPVCSAKIAARRKTDLQVVVDHAVKHGTMVSMLTLTQR 186
 QY 116 HTAQRLLHDTGLSAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGWHVHVHALL 175
 DB 187 HKQGLKHLMDALSTAWNRVTSGRRWIEFKEQFGLVGVYRANEITHGKNGWHVHVHALL 246
 QY 176 MFGSDVSENILESF-----SDAMFDRWTSKLVSLGFAAPLRNSGG 215
 DB 247 ISEKD---PLTSTFVYQKQGRRLPYPEIYMSDFIAERWEAGLAKHGVDV-LRDSGG 302
 QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVMEVSGDGKSGRHNRAPEIAVDA 266
 DB 303 LDWTVAKDARAIGN-----YVSKMOTSTDAISSEVTLGKFKKARNRTPFQILADI 354
 QY 267 VG-GDPQALELWREPEFGSMGRRRAIWSRGLRABAGLGAELTDQAIVQEESAPVMVAI 325
 DB 355 LSLGVDVDDLKWKYEKASFGRRALTWSKGLRDWANLGVQSDSEIA-SEIGDEAIALF 413
 QY 326 PARSWMMIRTCAPYVFG--EILGLVEAGA-----TWNLRDHLHYRLP 366
 DB 414 THDAWRQVRR-----FGAAELLDVTESGGRAAAAYRWLDPRE-IDWSLP 455
 RESULT 3
 ID AAW73071 standard; protein; 494 AA.
 AC AAW73071;
 XX
 XX 05-JAN-1999 (first entry)
 XX
 XX B. breve essential region gene protein.
 XX
 XX Essential region gene; shuttle vector.
 XX
 XX Bifidobacterium breve.
 XX
 XX JP10262670-A.
 XX
 XX 06-OCT-1998.
 XX
 XX 27-MAR-1997; 97JP-00091387.
 XX
 XX 27-MAR-1997; 97JP-00091387.

```

XX PA (HONS ) YAKULT HONSHA KK.
XX PN WPI; 1998-587288/50.
XX DR N-PSDB; AAV58945.
XX PD
XX PF Shuttle vector for a Bifidobacterium species - contains specific elements
XX PT from existing plasmids e.g. pNBb1 of ATCC 15698.
XX PR Claim 5; Page 14-15; 17pp; Japanese.
XX PS
XX PA This sequence is encoded by the essential region gene from plasmid pNBb1
XX CC used in the vector of the invention. The vector is a shuttle vector for a
XX CC Bifidobacterium, and contains: (a) a replicated essential region
XX CC originated from a plasmid pNBb1 of Bifidobacterium breve ATCC 15698; and
XX CC (c) an antibiotic-resistant gene originating from a plasmid of E. coli; and
XX CC an antibiotic-resistant gene originating from a plasmid of E. coli and
XX CC a shuttle vector can be used to transform various Bifidobacterium species
XX PT
XX PS Sequence 494 AA;
XX
Query Match 19.9%; Score 399.5; DB 2; Length 494;
Best Local Similarity 27.2%; Pred. No. 5.6e-31;
Matches 116; Conservative 64; Mismatches 169; Indels 77; Gaps 15;
QY 4 VSAEHLG-KDRP-----PVLVSSDKR-----GIRHELRPKLQITTTSETFN----- 44
Db 38 VAAEAASGARGSPPEWTKNKTIPSLRDLRLRAYGRRASRKILIVRHAGGETLGFEPK 97
QY 45 -----ACGRPISGVNGVTIVNGPKSGFGGLRSCGKGMICPCAGKVGHAHRADEISQVVA 99
Db 98 LPRCARGQPVD--TGUGVMINGEKARTGTMLCSIWACPTCSAIIIRHEAHEVALAIG 155
QY 100 HQL-----GTGSVAMVTMTMRHTAGQRLHDLMTGLSA 131
Db 156 NHAELKRAAADQQAHEGQRLPELMVSDSFGNVIPTLTLRHRTMTLAWTLDAIK 215
QY 132 AWKAATNGRRWRTRERMYGCGYVRAVEITGKNGWHVHVALMFSGDVSSENILESPSD 191
Db 216 GWTQVINGSPQWRASERWKIRGFVRAIBITYGVNGWHPHIFVMPFLDGLDDGQREAMQ 275
QY 192 AMEDRWTS--KLVLGPF-----AAPLRNSGLDVR-KIGGEADQVLAAYLTKIAGS 239
Db 276 WLLDRKWTWVRVAKAYKKDGNPNYVAP-NDEHGIDLFQKSGKDGAGTAAAYITKIQGD 334
QY 240 VG-----MEVSGDGKSGRHRNAPWEIYAVDAVGG--DPQALELWREFEFGSMGRRATW 292
Db 335 KGGVTLAQEIARGDIKNGRMGSVNPFLQLDSCGLSLDFQREDLWLEYQWQATLRRRCITW 394
QY 293 SRGLRARAGLAELTDAQIVQESAPVWVA-IIPARSWMMIRTCAPYVFGIILGLVEAG 351
Db 395 SRGLKEDMEV-EELEDELAEKADLPGLGVVFNRYVKDIRKSAPETTLADALDAER- 452
QY 352 ATWENL 357
Db 453 EDWQEV 458
RESULT 4
RAY97536
ID AAY97536 standard; protein; 427 AA.
AC AAY97536;
XX
XX DT 12-SEP-2003 (revised)
XX DT 15-JAN-2001 (first entry)
XX
XX DE B. lactofermentum p48K protein sequence.
XX
XX KW Temperature sensitive plasmid; TSRCR; protein production;
XX KW temperature sensitive replication control region; p48K.

```

```

OS Corynebacterium glutamicum.
XX
XX PN EP1038966-A1.
XX PD
XX PF 27-SEP-2000.
XX
XX PF 16-MAR-2000; 2000EP-00105326.
XX
XX PR 16-MAR-1999; 99JP-00069896.
XX
XX PA (AJIN ) AJINOMOTO CO INC.
XX
XX PI Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;
XX
XX DR WPI; 2000-573832/54.
XX
XX DR N-PSDB; AAA90934.
XX
XX PT Plasmids containing a temperature sensitive replication control regions
XX PT useful for breeding microorganisms for the production of amino acids by
XX PT fermentation.
XX
XX PS Claim 8; Page 14-17; 29pp; English.
XX
XX CC This sequence is the Brevibacterium lactofermentum p48K protein. The
XX CC invention relates to a plasmid containing a temperature sensitive
XX CC replication control region (TSRCR) and a marker gene (MG). The TSRCR is
XX CC derived from plasmid pAM330 harboured by Brevibacterium lactofermentum
XX CC (ATCC 13869) and allows the plasmid to replicate autonomously at a low
XX CC temperature but does not allow the plasmid to replicate autonomously at
XX CC an elevated temperature in coryneform bacteria within a temperature range
XX CC in which the bacteria can grow. The plasmid can be used for modifying a
XX CC chromosomal gene in a coryneform bacterium, which may be used for the
XX CC production of useful substances, such as amino acids, by fermentation, to
XX CC change their genetic traits. Therefore, the plasmid can be used for
XX CC breeding microorganisms for the production of amino acids by
XX CC fermentation. The plasmid comprises a TSRCR that allows the plasmid to
XX CC replicate autonomously at an elevated temperature in coryneform bacteria
XX CC within a temperature range in which the bacteria can grow. (Updated on 12
XX CC -SEP-2003 to standardise OS field)
XX
XX PS Sequence 427 AA;
XX
Query Match 16.3%; Score 326; DB 3; Length 427;
Best Local Similarity 29.1%; Pred. No. 1.2e-23;
Matches 98; Conservative 58; Mismatches 131; Indels 50; Gaps 15;
QY 34 LQOITTTSETFNACGR-----PISGVNGVTIVNGPKSGFGGLRSCGKGMICPCAGKV 86
Db 27 MYKITNSKALAGCHWRRRDEAVAVSWSS-----NG--ASQFEGQLNHSRWGSPLALELV 79
QY 87 GAHRADEISQVVAHQLTGTSVAMVTM-TMRHTAGQRLHDLMTGLSAKAAKATNGRRWRPTE 145
Db 80 MGERRIELATATKHLAAGGALMMFVGTVNRHSQSFAQVEAGIKTAYSSVVKTSQWKCE 139
QY 146 REMYCGDGYVRAVEITHG-KNGWHVHVALMFSGDVSSENILESPSDAMFDRWTSKLVS 204
Db 140 RARYGVSEHTYSDYEDTSDSWANGWHLHRNMLFLDRPLSDDELKAFESMFRWSAGVYKA 199
QY 205 GFAAPLRNSG-GLD-VRKIGGEADQVLAAYLTKIAGSGVMEVSGDGKSGRHRNAPWEI 262
Db 200 GMDAPLREHGKVLQDVSTWGGDAK-MATYLAKE--GMSQELTGSATKSGSYTTFQM 255
QY 263 AVDAVGGDPQALE-----LWREFEFGSMGRRAIARWGLRAGLAGAELTDAQI--- 311
Db 256 -LDMLAQSDAGEDMDAVLVARWREYEVGSKNLS-SWSRG--AKRALGIDYIDADVRE 311
QY 312 VEQE-----ESAPVMTVALIPARSWMMIRT 335
Db 312 MEELYKLAGEAPERVERSTVAVALKPDDMKLIQS 348
RESULT 5
ABB06341

```


Db 140 RARGVHTYSDYEVTSWANGHLHRNMLFLDRPLSDDELKAFESMFSRWAGVYKA 199
 Qy 205 GFAAPLRNSG-GLD-VRKIGEADQVLAAYLTKTASGVMEVSGDGKSGRHNAPWEI 262
 Db 200 GMDAPLREHGKLVQVSTWGGDAK-NATYLAK---GMSQELTGSATKTSKSGVTPFQM 255
 Qy 263 AVDAVGDPQALE-----LWREFEFGSGRAIAWSRGLRAGIGAEILTDAQI--- 311
 Db 256 -LDMLAQSDAGEDMDAVLVARWEYEVGSKNLR-SWSRG--AKRALGIDYIDAVRRE 311
 Qy 312 VEQE-----ESAPVMVAITPARSMMIRT 335
 Db 312 MEELYKLAGLEAPERVESTRVAVALKPDDWKLIQS 348

RESULT 7

ABB84283
 ID ABB84283 standard; protein; 451 AA.

AC ABB84283;

DT 06-AUG-2003 (revised)
 DT 13-JAN-2003 (first entry)

XX S. nigrificans pSN22 replication protein.

XX Plasmid stability protein; replication protein; ethylene forming enzyme;
 KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
 KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
 KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
 KW shuttle vector.

XX Streptomyces nigrescens.

XX WO200255709-A2.

XX 18-JUL-2002.

XX 12-DEC-2001; 2001WO-US047868.

XX 12-DEC-2000; 2000US-0254868P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Bramucci MG, Cheng Q, Kostichka KN, Tomb J;

XX WPI; 2002-557827/59.

XX New nucleic acid molecule encoding replication protein/plasmid stability
 PT protein, useful in cloning and expression vectors, particularly shuttle
 PT vectors for expression of heterologous genes in Rhodococcus species.

XX Example 5; Fig 4A; 96pp; English.

XX This invention describes a novel nucleic acid encoding a replication
 CC protein or a plasmid stability protein. The product of the invention is
 CC useful for expression of nucleic acid such as genes encoding enzymes
 CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
 CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
 CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
 CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
 CC Actinomycetales bacteria. The replication protein or plasmid stability
 CC protein are useful in cloning and expression vectors and particularly in
 CC shuttle vectors for the expression of homologous and heterologous genes
 CC in Rhodococcus sp. This sequence represents the S. nigrificans pSN22
 CC replication protein described in the disclosure of the invention.
 CC (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 451 AA;

Query Match 16.1%; Score 322; DB 5; Length 451;
 Best Local Similarity 30.8%; Pred. No. 3.2e-23;

Matches 101; Conservative 41; Mismatches 110; Indels 76; Gaps 15;
 Qy 49 PISGV-----NGVTIYNGPKSGFGILRSCGKWI CPCACGKYGARADHSIQVVAHQL 102
 Db 3 PASGVI VAQTAAAGTSVVL-----GLMRCGR IWLCPVCAATIRHKRAEBITAAVVVEWI 54
 Qy 103 GTGSVA-MVTMTMHTAGQRHLDLWTGLSAAWKAATNGRR-----WRTE--- 145
 Db 55 KRGGTAYLVITFTARHGHDTDLADLMDALQGRKTDADAPRRPGAYQRLITGGTWAGRRAKD 114
 Qy 146 -----REMYGCDGYVRAVEITHGK-NGMHVHVHLMFSGDV-----SENILESF 189
 Db 115 GHRAADREGIRDRIGYVGMIRATEVTVQINGWHPHIAIIVLGVGRTEGERSAQIVGTF 174
 Qy 190 --SDAMFDRWTSKLVSLGFAPIR-----NSGGLDVRKIGIGEAD-QVLAAYLTQIA 237
 Db 175 EPSEALDEWQGWRAV-WTAALKRVNPFQTPDDRHGVDKFKLETERDANDLABYIAKTQ 233
 Qy 238 SG--VGMEVSGDGKSGRHNAPWEI--AVDAVG--DPQA-----LELWREFE 281
 Db 234 DGKAPALELARADLKTANGNVAFELLGRIGDLTGWTEDDAAGVGSLEWNLARWHEYE 293
 Qy 282 FGSNGRAIAWSRGLRAGIGAEILTDA 309
 Db 294 RATKGRRAIEWTRYLRQMLGLDGGDTEA 321

RESULT 8

AAAY97537

ID AAAY97537 standard; protein; 427 AA.

XX AAAY97537;

DT 12-SEP-2003 (revised)

DT 15-JAN-2001 (first entry)

XX B. lactofermentum pSFK6 protein sequence.

XX Temperature sensitive plasmid; TSRCR; protein production;
 KW temperature sensitive replication control region; pSFK6.

XX Corynebacterium glutamicum.

XX EP1038966-A1.

XX 27-SEP-2000.

XX 16-MAR-2000; 2000EP-00105326.

XX 16-MAR-1999; 99JP-00069896.

XX (AJIN) AJINOMOTO CO INC.

XX Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;

XX WPI; 2000-573832/54.

XX N-FSDB; AAA90951.

XX Plasmids containing a temperature sensitive replication control regions
 PT useful for breeding microorganisms for the production of amino acids by
 PT fermentation.

XX Disclosure; Page 19-22; 29pp; English.

XX This sequence is the Brevibacterium lactofermentum pSFK6 protein. The
 CC invention relates to a plasmid containing a temperature sensitive
 CC replication control region (TSRCR) and a marker gene (MG). The TSRCR is
 CC derived from plasmid pAM330 harboured by Brevibacterium lactofermentum
 CC (ATCC 13869) and allows the plasmid to replicate autonomously at a low
 CC temperature but does not allow the plasmid to replicate autonomously at
 CC an elevated temperature in coryneform bacteria within a temperature range
 CC in which the bacteria can grow. The plasmid can be used for modifying a
 CC chromosomal gene in a coryneform bacterium, which may be used for the

XX Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW shuttle vector.
XX Streptomyces lividans.
OS WO200255709-A2.
FN PD 18-JUL-2002.
XX PF 12-DEC-2001; 2001WO-US047868.
XX PR 12-DEC-2000; 2000US-0254868P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX WPI; 2002-557827/59.
XX New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.
XX Example 5; Fig 4A; 96pp; English.
XX This invention describes a novel nucleic acid encoding a replication
CC protein or a plasmid stability protein. The product of the invention is
CC useful for expression of nucleic acid such as genes encoding enzymes
CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence represents a Rhodococcus AN12 Rep
CC protein described in the disclosure of the invention
XX Sequence 456 AA;
Query Match 15.8%; Score 316; DB 5; Length 456;
Best Local Similarity 30.5%; Pred. No. 1.3e-22;
Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;
QY 49 PISGV-----NGVTIVNGPKSGFGGLRSCKGKWCPCCKAGVGAHRADEISQVVAHQ 102
DB 3 PASGVIVAQTAAAGTSVVL-----GLMRCGRILWCPVCAATIRKRAEITAAVVEWI 54
QY 103 GTGSA-VYTMTRHTAGORLHDLWTGLSAAMKAATNGRR-----WRTE----- 145
DB 55 KRGGTAYLVTFARHGHTDRDLADLMDALQGRKTPDSPRRPGAYQRLTGTGWAGRRAND 114
QY 146 -----REMYCDGVVRAVEITHGK-NGWVHVHALLMFSGDV-----SENLSEF 189
DB 115 GHRAADRIGRIDRIGYGMIRATEVTVGQINGWPHIHAIVLVGRTGERSAKQIVATF 174
QY 190 --SDAMFRWTSKLVSLGFAAPLR-----NSGGLDVRKIGGEAD-QVLAAYLTAKIA 237
DB 175 EPTGAALDEWQHWSV-WTAALRKVNPAFTDDRHGVDVFKLETERDANDLAEYIAKTQ 233
QY 238 SG--VGMVEGSDGSGRHNAPWEI---AVDVG-----DPOA-----LELWREFE 281
DB 234 DGKAPALEIARADLTATGNNVAPPELLGRIGDLTGGMTEDDAAGVGSLEWNLRSWHEYE 293
QY 282 FSGMRRATAWSRGLRARAGLGAELTDA 309
DB 294 RATRGRRALEWTRYLRQMLGLDGGDTEA 321

RESULT 13
AAU11039
ID AAU11039 standard; protein; 466 AA.
XX AC AAU11039;
XX 29-AUG-2003 (revised)
DT 07-AUG-2003 (revised)
DT 12-MAR-2002 (first entry)
XX DE
XX Replication protein encoded by Ketogulonigenium plasmid pADMX6L2.
KW Ketogulonigenium; 2-keto-L-gulononic acid; L-sorbose; sorbitol;
KW replication protein; plasmid pADMX6L2.
XX Ketogulonigenium robustum.
XX WO200177348-A2.
XX 18-OCT-2001.
XX PF 05-APR-2001; 2001WO-US011097.
XX PR 05-APR-2000; 2000US-0194627P.
XX PA (ARCH) ARCHER-DANIELS MIDLAND CO.
PA (DELI/) D'ELIA J.
PA (STOD/) STODDARD S F.
XX D'elia J, Stoddard SF;
XX WPI; 2002-041295/05.
DR N-PSDB; AAS17120.
XX New bacterium of Ketogulonigenium genus, useful for producing 2-keto-L-
PT gulonic acid from sorbose or sorbitol, comprises transgene containing DNA
PT sequence from endogenous Ketogulonigenium plasmid.
XX Disclosure; Fig 9; 116pp; English.
XX The present invention relates to a new bacterium of genus
CC Ketogulonigenium. Ketogulonigenium may further comprise a transgene,
CC comprising a DNA sequence from an endogenous Ketogulonigenium plasmid.
CC Methods for transforming Ketogulonigenium are also described. The
CC invention is useful for producing 2-keto-L-gulononic acid (2-KLG) from L-
CC sorbose or sorbitol. The present sequence represents the replication
CC protein encoded by the Ketogulonigenium endogenous plasmid pADMX6L2.
CC (Updated on 07-AUG-2003 to correct OS field.) (Updated on 29-AUG-2003 to
CC standardise OS field)
XX Sequence 466 AA;
Query Match 14.1%; Score 282; DB 5; Length 466;
Best Local Similarity 30.0%; Pred. No. 3.5e-19;
Matches 87; Conservative 40; Mismatches 119; Indels 44; Gaps 15;
QY 60 NGPKSGFGGLRSCKGKWCPCCKAGVGAHRADEISQVVAHQLGTSV-AMVTMTMRHTA 118
DB 109 DGGRAFFSGGLAQCRNVMGCAVCSARIAQIRSEMHLLAWARDNGFVPLVLTIAQHK 168
QY 119 GORLHDLWTGLSAAMKAATNGRRWRTEREMYGCD-----GYVRAVEITHG-KNGMWHVHA 173
DB 169 GDSLFDLLQNMKKAQRLRQRREW-----DLFPVGSVTSTSTIHSYANGWHPHFHE 220
QY 174 L-LMFSGDVSE--NILESFSAMFDRWTSKLVSLGFAAPLRNSGGLDVRKIGGEADQVLA 230
DB 221 IVLLRAGDESEALHLMQRLGDA---WRACLKGYGMWG---NDAADFVRGAANAGD---- 269
QY 231 AYLTKIASGVGMVEGSGDKSGRHNAPWEIADVAGDPCQALBELWREFEFGSMG--RR 288
DB 270 -YVAK--WGAABELTSSSKGKRKRTPRQLL--QAGDD-----GLWLEYFNATSGKRRR 320
QY 289 AIAWSRGLRARAGLGAELTD-----AQIVEQESAFVMVAIIPARSMIMIR 334

SQ Sequence 314 AA;

Query Match	6.7%;	Score 134;	DB 1;	Length 314;
Best local Similarity	22.4%;	Pred. No. 0.0001;		
Matches	56;	Conservative	50;	Mismatches 106;
			Indels	38;
			Gaps	10;
Qy	75	KGWICPCAGKVGAAHRADEISQVV--AH-QLGTSVAMVTMTMRHTAGQRIHDLWTGLSA	131	
Db	69	KSRLCPLCNWRRSMGQSNQLMQVLDEAHKQKTKGRFLFLTLTAENASGLENLKQEVKRWGR	128	
Qy	132	AWKAATNGRWRTREMYGCDGYRAVEITHKNG-MHVHVHALLMESGDVSENILSFSS	190	
Db	129	A-----ISKLFQYKKPKAKNLLGYVRSTEITINKNGTHQGMHVLLFKVPYFKDSANYIN	183	
Qy	191	DAMPDRWTSKLIVSLGFAPLRNSGGLDVRIKGGADQVLAAYLTKIASGVGMVSGDGGK	250	
Db	184	DAEWSKULXRAMKLDY-XPIVNVAVRSNKAAGKNSLIASAQET-----	226	
Qy	251	SGRHGNAPWEI-AVDVAGGDPQA-LELWREFFEGSMGRRAIAWSRGL--RARAGLGAEI	306	
Db	227	-----AKYQVKSKDILTDQERDLQVVEDEQLAGSROIYSG-GLFKFIRKQLQLED	278	
Qy	307	TDQIVQEQQE	316	
Db	279	VDAMLINVD	288	

Search completed: October 23, 2004, 03:55:45
Job time : 88 secs

Db 312 MEELYKLAGLEAPERVESTRAVALVKPDDMKLIQS 348

RESULT 2

US-09-521-668B-20
; Sequence 20, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: MATSUI, KAUKIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521.668B
; CURRENT FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: JP 11-69896
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-09-521-668B-20

Query Match 15.9%; Score 318; DB 3; Length 427;
Best Local Similarity 28.8%; Pred. No. 3.7e-24;
Matches 97; Conservative 58; Mismatches 132; Indels 50; Gaps 15;
QY 34 LQOITTSFTNACGR-----PISGVNGVTIVNGPKSGRGLRSCGKGMICPCCAK 86
DB 27 MYKITNSKALAGCHRRRDEAVASWS--NG--ASQFEGJONSHSRMGSSLAEL 79
QY 87 GAHRADEISQVVAHQLTGTSVAMVTM-TMRHTAGQRLHDLMTGLSAAKAAATNGRRWTE 145
DB 80 MGERRIELAIATKHLAAGGALMEVGIVRNRSSQFPAQVAGIKTAYSSVKTQSQWKE 139
QY 146 REMTGCDCGVRAVEITHG-KNGWHVHALLMFSGDVSENILESPSDAMFRTWSKL 204
DB 140 RARYGVHEITYSDYETDWSWANGHLHRNMLFLDRLPSDDLKAFEDSMFSRWAGVYKA 199
QY 205 GFAAPLRNSG-GLD-VRKIGEAQOVLAAYLTKTASGVGMVSGDGKSGRHRNAPWEI 262
DB 200 GMDAPLRHGVKLDQVSTWGGDAK-MATYIAK---GMSQELTGSATYATASKGSTYTPQM 255
QY 263 AVDAVGDDPOALE-----LWREFEFGSMGRRRAIWSRGLRAGLAGELTDAQI--- 311
DB 256 -LDMLAQSDAGEDMDAVLVARWREYEVGSKNLR-SWSRG--AKRALGIDYIDADVRE 311
QY 312 VEQE-----ESAPVMAIIPARSWMIRT 335
DB 312 MEELYKLAGLEAPERVESTRAVALVKPDDMKLIQS 348

RESULT 3

US-09-252-991A-17653
; Sequence 17653, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17653

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (730)

; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

US-09-252-991A-17653

Query Match 5.1%; Score 101.5; DB 4; Length 882;
Best Local Similarity 21.7%; Pred. No. 0.26;
Matches 87; Conservative 48; Mismatches 125; Indels 141; Gaps 20;
QY 47 GRPISGVNGVTIVNGPKSGRGLRSCGKGMICPCA--GKVGAAHRADEISQVV----- 98
DB 487 GRVLGGVIGLALRQCAIGNR-----VAPFVAQAGRKVGLEAAEQLSRPMHRPTE 535
QY 99 -----AHQLGTGS-----VAMVTMTMRHTAGQRLHDL-----W 126
DB 536 IRHAPADGVGGGAGRRARELVGGGGQQLRLLVVQVRIAL-HDQQRAGGVGRHGTGL 594
QY 127 TGLSAAWKAATNGRRWRTEREMYGCDGVRA-----VEITH-----GKNGHV 169
DB 595 VGVAAAGHGAVDQAAGRGDAFVLGDAAAVVALAVLLVEAGHGQPVAFQVRLEVGQGAHA 654
QY 170 HV-----HALLMFSGDVSENILESPSDAMFD---RWTSKLVS--LGFAAPL- 210
DB 655 GVGVAAGAEVDVHAL---AGDAGGVQVPAGAPVILGIGGVAVEGLVADVGLUAATV 711
QY 211 -----RNSGGLDVRKIGTGGEADQVLAAYLTKTASGVGMVSGDGKSGRHRNAPW 260
DB 712 VDRPHPGVQRLVGGLEIXRVRRAREQAVVL-----VGVDVHLRCRVGH--- 756
QY 261 EIAVDVAG-GDPOALEW-----REPEFGSMGR-RATAWRSGLR-ARAG 301
DB 757 PVHADAVARGAAGAADVGAGVIVAVHRAADAERRAVIDGAAGRHLVARGGGIRMARVE 816
QY 302 LGAELTDAQIIVEQESAPVMAIIPARSWMIRTCAPVVF 342
DB 817 TGVHLADLHALAVDAGGIGLVGPHFPAPVLL-----VFG 851

RESULT 4

US-09-252-991A-30345
; Sequence 30345, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30345
; LENGTH: 1395
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30345

Query Match 5.0%; Score 101; DB 4; Length 1395;
Best Local Similarity 24.5%; Pred. No. 0.57;
Matches 104; Conservative 34; Mismatches 162; Indels 124; Gaps 24;
QY 13 DRPVLVSSDKRGIRHRLRPKLQITTTSETNACGR--PIGQ----VNGVTIVNGPKSG 66
DB 511 DRIAVAGDEHRAAR--FRSAEAAAQVQDGRGRRAIDGGAALAAAGVARRIGRGVD 568
QY 67 FGGLRS-CGKGWICPCCAKGVGAHRADEISQVV-----AHQIG-----TGSVAMVTWTRH 116

Db 569 HGSIRQRRGRG-EAPVAAA-VGGDLADRVAVAIQGHAAARLGGAAERGAAGVVD-----H 622
QY 117 TAGQ--RLHDLWTGLSAWKAATNG-----RWRTEREMVGC-----DGYV 155
Db 623 WRGRGYRVH-----GDTAARTGAVGGVGRGVDHGAAGVQRRGRSEAPVAAAIGGDLPGIA 678
QY 156 RAVEITHKNGHVVHALLMFSGDVSENILESFSDAMFDRWTSKLVSLGFAAPLR----- 211
Db 679 VAICQGHG-----CARLGGAERGAARVAVDHWGRCHGVHGDAGRRTGVA 723
QY 212 ---NSGGLDVKKIG-----GEADQVLAAYLTKIASGVMEVSGDGKSGRHNAPWEIA 263
Db 724 GRVSGRGVDHRAVQRRARGEG-PVAAAAGVDLANRVAVAVQGHG-SARLGGAEGRTV 781
QY 264 VDAVGDPQALELWREFSGMGR-AIAW-----SRGLR-----ARAGLGA 305
Db 782 ARIDGGRG-----RDGHNAGRTAVAGGIPGHVHDHGTVRQRRGRERPAVVGIGG 836
QY 306 LTDAQIVEQEBESAPVMVAITPARSMMIRTCAPVVFGEIILGLVEAGATWENLRDLHLVRL 365
Db 837 LAD-----RIAIVGQGHAAARFGSPAFAVARL-----DHRVR 873
QY 366 PAAD 369
Db 874 HAVD 877

RESULT 5

US-09-252-991A-23181
; Sequence 23181, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23181
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23181

Query Match 5.0%; Score 99.5; DB 4; Length 676;
Best Local Similarity 25.6%; Pred. No. 0.28;
Matches 89; Conservative 31; Mismatches 135; Indels 93; Gaps 19;
QY 62 PKSGFGGLRSCGKGMIC--PCCAGKVGHAHRADEIS-----QVVAHQLTGSAVMVTMT 113
Db 318 PRGPGAAGRAGSGCGCGCGAGRTAGGASPVATGSGDCCSRAPAPVA----- 372
QY 114 MRHTACORLHDLWTGLSAWKAATNGRWRMTEREMYGCDGYVPAVRITHGKNGHVVHVA 173
Db 373 -GRGAGNRRRRRWSLPAS--CLSRGR-----GVCRSVARRPG----- 409
QY 174 LMFSGDVSENILESFSDAMFDRWTSKLVSLGFAAPLRNSGGLDVVKIGGEADQVLAAYL 233
Db 410 -AVRRGSVCALLTOAEKAAGD-----AVDLGLGA-----YRLDAQRFGENARQFLAACR 458
QY 234 TKIASGVG-MEVGS-----GDGK-----SGRHGNRAPWEIAYD--AVGSD-POALELWRE 279
Db 459 GKERRAFGETEGALFQAHEGRQLRDLVRQAGPAQFGFVVDHQAGGVDQAOLE----- 514
QY 280 FEPGSGMR--RAJAWSRGLRARAGLGAELTDAQIVEQESAPV--MVAIIPARSMWIR 334
Db 515 -----GRMVQATEQTEGVAE-----QGDVGEAQVQAQRGPAHRSVAVVDDAQSHLLA 564

QY 335 TCAPYV-----FGEILGLVEAGA--TWNELRDHLHYRLPAAD 369
Db 565 GRAALVHLEEDHVPPRHDEGQQQAEGAGAEAEERQAEHOFHQDQAAD 612

RESULT 6

US-09-252-991A-19231
; Sequence 19231, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19231
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19231

Query Match 4.9%; Score 98.5; DB 4; Length 621;
Best Local Similarity 22.4%; Pred. No. 0.31;
Matches 62; Conservative 27; Mismatches 87; Indels 101; Gaps 11;
QY 175 LMFSGDVSENILESFSDAMFDRW-----TSKLVSLGFAAPLRNSGGL 216
Db 323 LLVSGVDNSLLGSLYDLPQRFHLCLGDEESLLPHRLQRFELRQEAFFMLLRAVG- 381
QY 217 DVKIGGEADQVLAAYLTKIASGVME-----VCSGDKSGRHNAPWEIAYDVGDDPQ 272
Db 382 ---NFGGATRMSSLLMYQLRADGIGEGYHCVLLGEG-----ADELFMWGYP 425
QY 273 ALELWREFFGSMGRRRAIARAGIARAGIAGLTAQIVEQ-EESA----- 318
Db 426 HLELWRRRDAPERPFAAAWFGYRKAALLAEPAGRRVAERIEELAEALGQGLEAAIG 485
QY 319 --PVMVAIIP-----ARSWMMIRTC--APYVFG-----EILG- 349
Db 486 QFDLHYSLEPLLRADHLLMSRTIEARTPYLHGALAQACRGLQRIQVGD TAKAPIVALLEQ 545
QY 350 AGATWE-----NLRDHLHYRLPA 367
Db 546 REKRWQAQPKRHRLPFRFPWPQALGEMRRHLAERLPA 582

RESULT 7

US-09-252-991A-22282
; Sequence 22282, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22282
; LENGTH: 1409
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22282	
Query Match 4.9%; Score 98.5; DB 4; Length 1409;	
Best Local Similarity 23.8%; Pred. No. 1;	
Matches 103; Conservative 51; Mismatches 160; Indels 119; Gaps 24;	
Qy	14 RPPVLSSDKGRHRLR--PKLQIITSETFNACGRPISGVNGVTI-----VNGP 62
Db	275 RPFALDAQSGGILQVQRCPGLQLSR-----RLAGRDLAGPAGETVDRPRDPLHGLGY 330
Qy	63 KSGS-----FGGLRSCKGWIC--PCCAG-----KVGARADEISQVVAHQ 101
Db	331 RQSGADHPVAGCALLEEIEGQLQAAIPEVVAAADPHLAALQORQAGARAETIQCAQLQ 390
Qy	102 L-----GT-----GSVAMVTWTRHTAGOR---LHDLWTGLSAANKAATN----- 138
Db	391 LAIAGLRQPGTGVQGDVAAATAQSLAVGQQRPRGHVEQTALSOAQVAAAAGNPSAVAV 450
Qy	139 GRR---WTEREMVCGDYVRAVEITHGKNGWHVHALLMFSGDVSENILESFSDAMFD 195
Db	451 GQRIVLRRQQQVAPQAAGIOPVAVAHQAQA-----VFAAGIDETI-----DRPVR 496
Qy	196 RWTSKVLSLGFAPLRNSGGLDVRKIGGEADQVLAAYLTKIASGVGMEVSGDKSGRHG 255
Db	497 RNLQHVAGIQAAMGSDGVGD-RHVVGEDQ-CLAAI-----GQQRADVADALAGRHL 548
Qy	256 NRAPWEIADAVGGDPQ---ALELWREFEFGSMGRRRAIAWSRGLRABAGLGAELTDAQIV 312
Db	549 QAAQLE-AVEQFAVQAQAALAGGHRSGVQGRLLGRRLAGTDHLLLEAAGIGAE-----SVL 603
Qy	313 EQESAPVMVAIIP-----ARSMWMTTCAPYVFGELGIVEAGATWE-----NLR 358
Db	604 QQAALHVGCVGAPDDQLCAARHGIL-----RLVAVQAG-LWRFGLPAAVGVA 652
Qy	359 DHL--HYRLPAAD 369
Db	653 DHLAVEPRIGNAD 665
RESULT 8	
US-09-252-991A-30935	
; Sequence 30935, Application US/09252991A	
; Patent No. 6551795	
; GENERAL INFORMATION:	
; APPLICANT: Marc J. Rubenfield et al.	
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS	
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS	
; FILE REFERENCE: 107196.136	
; CURRENT APPLICATION NUMBER: US/09/252,991A	
; CURRENT FILING DATE: 1999-02-18	
; PRIOR APPLICATION NUMBER: US 60/074,788	
; PRIOR FILING DATE: 1998-02-18	
; PRIOR APPLICATION NUMBER: US 60/094,190	
; PRIOR FILING DATE: 1998-07-27	
; NUMBER OF SEQ ID NOS: 33142	
; SEQ ID NO 30935	
; LENGTH: 1245	
; TYPE: PRT	
; ORGANISM: Pseudomonas aeruginosa	
US-09-252-991A-30935	
Query Match 4.9%; Score 97.5; DB 4; Length 1245;	
Best Local Similarity 21.2%; Pred. No. 1.1;	
Matches 86; Conservative 50; Mismatches 143; Indels 127; Gaps 19;	
Qy	9 LSGKDRP--PVLVSSDKRGIRHELKPKLQIITSETFNACGRPI-----SGVNGVT 57
Db	188 LRGQPSGPAFPGRQRPQGRHTRYVQO-----RPAQAGPRRGGETGARGVA 235
Qy	58 IVNGPKSGFGGLRSCGWG-----ICPCCAKVGAHRADEISQVVAHQ 102
Db	236 AVPGE-----GCTCAAGWRQAPGPRRAAPGTRPCGAGAGQLARWATGVPRQAAGRT 289
US-09-335-409-11	
; Sequence 11, Application US/09335409	
; Patent No. 6121029	
; GENERAL INFORMATION:	
; APPLICANT: Schupp, Thomas	
; APPLICANT: Ligon, James	
; APPLICANT: Molnar, Istvan	
; APPLICANT: Zirkle, Ross	
; APPLICANT: Cyr, Devon	
; APPLICANT: Goerlach, Joern	
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES	
; FILE REFERENCE: 4-30582A	
; CURRENT APPLICATION NUMBER: US/09/335,409	
; CURRENT FILING DATE: 1999-06-17	
; NUMBER OF SEQ ID NOS: 30	
; SOFTWARE: PatentIn Ver. 2.0	
; SEQ ID NO 11	
; LENGTH: 713	
; TYPE: PRT	
; ORGANISM: Sorangium cellulosum	
US-09-335-409-11	
Query Match 4.8%; Score 97; DB 3; Length 713;	
Best Local Similarity 22.1%; Pred. No. 0.55; Indels 80; Gaps 12;	
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;	
Qy	30 LRPKLQIITSETFNACGRPISGVNGVTIIVNGPKSGFGGLRSCGKMWICCCAGKVGAAH 89
Db	302 MRVDVSLRTPAAGTVALLLATATAAKVVPALGARLGLRGSEALVAVGLNWKGG-- 359
Qy	90 RADEISQVVAHQIGTGS-----VAMVTWTRHTAGORLHDLWTGLSAANKAATN 140
Db	360 -TDLIVAIVGVVELGLLSNEAYTMVAVVALTV---TASPAALL-IMLEKRA---PPTQEE 410
Qy	141 RWRTEREMVCGDYVRAVEITHGKNGWHVHALLMFSGDVSENILESFSDAMFDRWTSK 200
Db	411 SARLEREAAARRAYIPGVE---RILVPIVAHUPGFATDIVESIVAS----- 454
Qy	201 LVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYLTKIASGVGMEVSGDKSGRHNGR--- 257
Db	455 -----KRKLGETVD-ITELSVQQAQPGPSRAAGEASRGLARLARLV 496
Qy	258 APWEIADVAGDQPALELWREFEFGSMGRRRAIAWSRGLRABAGLGAELTDAQIVEQES 317
Db	497 GIMQRRELGRSIOAILRASRDHLLVIGARSPARAGMSF-----GRLDQAIVQRAESN 551
Qy	318 APVMVAIIPARSWMIRTCPYVFGELGIVEAGATWENLDRHLHYRLPAADV 370
Db	552 VLVVVGDPFAAE---RASARRILVPIG-----LEYSFAAADL 586

```
RESULT 10
US-09-568-102-11
; Sequence 11, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-568-102-11

Query Match      4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

QY 30 LRPKLOQITSETFNACGRPISGVGTIYVNGPKSGFGGLRSCGKWCIPCCAGKVGAAH 89
Db 302 MRVDVSQRLTPAAWGTVALLLATATAAKVVPAAALGARLGLRGSEAAALVAVGLNMKG-- 359
QY 90 RADEISQVVAHQLTGS-----VAMVTMTMRHTAGORLHDLWTGLSAAMKAATNGR 140
Db 360 -TDLIVAIVGVVELGILLSNEAYTMVAVVALTV-----TASPALL-IWLEKRA---PPTQEE 410
QY 141 RWTEREMVCGDYVRAVEITHGKNGWHVHALLMFSGDVSENILESFSDFAMFDRWTSK 200
Db 411 SARLEEEAARRAYIPGVE-----RILVPIVAHALPGFATDIVESIVAS----- 454
QY 201 LVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYITKIASGVGMVSGDGKSGRHN--- 257
Db 455 -----KRKLGETVD-ITELSEVQAPGPSRAAGEASRGLARLGLRLV 496
QY 258 APWEIADVAGDPPQALELWREFEFGSMGRRAIAWSRGLRARAGLGAELTDAQIVEQEE 317
Db 497 GIWRQRRELRGSIQAILRASRDHLLVIGARSPARARGMSF-----GRLODAIVQRAESN 551
QY 318 APVWVAIIPARSMWMTIRTCAPYVFGELGLVEAGATWENLRDLHYRLPAADV 370
Db 552 VLWVVGDPAPAE-----RASARRILVPIIG-----LEYSFAAADL 586

RESULT 11
US-09-567-969-11
; Sequence 11, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 11
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-567-969-11

Query Match      4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

QY 30 LRPKLOQITSETFNACGRPISGVGTIYVNGPKSGFGGLRSCGKWCIPCCAGKVGAAH 89
Db 302 MRVDVSQRLTPAAWGTVALLLATATAAKVVPAAALGARLGLRGSEAAALVAVGLNMKG-- 359
QY 90 RADEISQVVAHQLTGS-----VAMVTMTMRHTAGORLHDLWTGLSAAMKAATNGR 140
Db 360 -TDLIVAIVGVVELGILLSNEAYTMVAVVALTV-----TASPALL-IWLEKRA---PPTQEE 410
QY 141 RWTEREMVCGDYVRAVEITHGKNGWHVHALLMFSGDVSENILESFSDFAMFDRWTSK 200
Db 411 SARLEEEAARRAYIPGVE-----RILVPIVAHALPGFATDIVESIVAS----- 454
QY 201 LVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYITKIASGVGMVSGDGKSGRHN--- 257
Db 455 -----KRKLGETVD-ITELSEVQAPGPSRAAGEASRGLARLGLRLV 496
QY 258 APWEIADVAGDPPQALELWREFEFGSMGRRAIAWSRGLRARAGLGAELTDAQIVEQEE 317
Db 497 GIWRQRRELRGSIQAILRASRDHLLVIGARSPARARGMSF-----GRLODAIVQRAESN 551
QY 318 APVWVAIIPARSMWMTIRTCAPYVFGELGLVEAGATWENLRDLHYRLPAADV 370
Db 552 VLWVVGDPAPAE-----RASARRILVPIIG-----LEYSFAAADL 586

RESULT 12
US-09-568-480-11
; Sequence 11, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-568-480-11

Query Match      4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

QY 30 LRPKLOQITSETFNACGRPISGVGTIYVNGPKSGFGGLRSCGKWCIPCCAGKVGAAH 89
Db 302 MRVDVSQRLTPAAWGTVALLLATATAAKVVPAAALGARLGLRGSEAAALVAVGLNMKG-- 359
QY 90 RADEISQVVAHQLTGS-----VAMVTMTMRHTAGORLHDLWTGLSAAMKAATNGR 140
Db 360 -TDLIVAIVGVVELGILLSNEAYTMVAVVALTV-----TASPALL-IWLEKRA---PPTQEE 410
QY 141 RWTEREMVCGDYVRAVEITHGKNGWHVHALLMFSGDVSENILESFSDFAMFDRWTSK 200
```

Patent No. 6358719
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,472
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 713
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-472-11

Query Match 4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

Qy 30 LRPKLOQITSETFNACGRPIGNGVTVVNGPKSGFGGLRSCGKGMWICPCCKAGKVGAAH 89
Db 302 MRVDVSQLRTPAAGTVALLLATATAAKVVPALGALGLRGSEAAALVAVGLNNKGG-- 359
Qy 90 RADEISOVVAHQIGTGS-----VAMVTMTMRHTAGORLHDLTGLSAWKAATNGR 140
Db 360 -TDLIVAIVGVELGLLSNEAYTMVAVVAVTV---TASPALL-IWLEKRA---PPTQEE 410
Qy 141 RWRTERMYGCDGVVRAVEITHGKNGWHVHALLMFSDVSENILSFSDAMFDRWTSK 200
Db 411 SARLEREEAARRAIPGVE-----RILVPIVAHALPGFATDIVESIVAS----- 454
Qy 201 LVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYLTAKIASGVGMVSGDGKSGRHN-- 257
Db 455 -----KRKLGETVD--ITELSVQQAPGPSRAAGASRGLARGLARLV 496
Qy 258 APWEIADVAGDQPALELMREFEFGSMGRRAIAWSRGLRARAGLGAELTDAQIVQEE 317
Db 497 GIWQRRELRSIQAILRASRDHLLVIGARSPARAGMSF-----GRLODAIVQRAESN 551
Qy 318 APVVAIIPARSMWMTCPAPYVGEILGLVEAGATWENLRDHLHYRLPAADV 370
Db 552 VLVVVGDPAAE-----RASARRILVPIIG-----LEYSFAAADL 586

RESULT 15
US-09-567-899-11
Sequence 11, Application US/09567899
Patent No. 6383787
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/567,899
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 713
TYPE: PRT

Patent No. 6358719
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,472
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 713
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-472-11

Query Match 4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

Qy 30 LRPKLOQITSETFNACGRPIGNGVTVVNGPKSGFGGLRSCGKGMWICPCCKAGKVGAAH 89
Db 302 MRVDVSQLRTPAAGTVALLLATATAAKVVPALGALGLRGSEAAALVAVGLNNKGG-- 359
Qy 90 RADEISOVVAHQIGTGS-----VAMVTMTMRHTAGORLHDLTGLSAWKAATNGR 140
Db 360 -TDLIVAIVGVELGLLSNEAYTMVAVVAVTV---TASPALL-IWLEKRA---PPTQEE 410
Qy 141 RWRTERMYGCDGVVRAVEITHGKNGWHVHALLMFSDVSENILSFSDAMFDRWTSK 200
Db 411 SARLEREEAARRAIPGVE-----RILVPIVAHALPGFATDIVESIVAS----- 454
Qy 201 LVSLGFAAPLRNSGGLDVRKIGGEADQVLAAYLTAKIASGVGMVSGDGKSGRHN-- 257
Db 455 -----KRKLGETVD--ITELSVQQAPGPSRAAGASRGLARGLARLV 496
Qy 258 APWEIADVAGDQPALELMREFEFGSMGRRAIAWSRGLRARAGLGAELTDAQIVQEE 317
Db 497 GIWQRRELRSIQAILRASRDHLLVIGARSPARAGMSF-----GRLODAIVQRAESN 551
Qy 318 APVVAIIPARSMWMTCPAPYVGEILGLVEAGATWENLRDHLHYRLPAADV 370
Db 552 VLVVVGDPAAE-----RASARRILVPIIG-----LEYSFAAADL 586

RESULT 15
US-09-567-899-11
Sequence 11, Application US/09567899
Patent No. 6383787
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/567,899
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 713
TYPE: PRT

```
; ORGANISM: Sorangium cellulosum
US-09-567-899-11

Query Match      4.8%; Score 97; DB 3; Length 713;
Best Local Similarity 22.1%; Pred. No. 0.55;
Matches 78; Conservative 38; Mismatches 157; Indels 80; Gaps 12;

QY 30 LRPKLQIITSETNACGRPISGVNGVTIVNGPKSGFGGILRSCGKWCIPCCAGKVGAAH 89
Db 302 MRVDVSQIRTPAAAGTVALLIATATAAKVPAALGARLGGILRGSEAAALVAVGLNMKG-- 359

QY 90 RADEISQVVAHQLTGTS-----VAMVTMTMRHTAGQRLHDLWTGLSAAWKAAATNGR 140
Db 360 -TDLIVAIVGVELGSLNEAYTMVAVVALTV----TASPALL-IWLEKRA---PPTQEE 410

QY 141 RWRTERMYGCDGYVRAVEITHGKNGMHHVHALLMFGSDVSENILESFDAMFDRWTSK 200
Db 411 SARLEREEAARRAYIPGVE----RILVPIVAHALPGFATDIVESIVAS----- 454

QY 201 LVSLGFAAPLRNLSGLDVRKIGGEADQVLAAYLTKIASGVGMEVSGDGKSGRHCNR--- 257
Db 455 -----KRKLGETVD--ITELSVQQAPGPSRAAGEASRGLARLGARLRY 496

QY 258 APWEIAVDVAGGDPQALELWREFEFGSMGRRATAWSRGLRARAGLGAELTDAQIVEQEE 317
Db 497 GIWQRRELGRSIGAILRASRDHDLVLVIGARSPARARGMSF-----CRLQDAIVQRAESN 551

QY 318 APVNVAILPARSWMINTCAPYVFEIGLVEAGATWENLRDHLHYELPAADV 370
Db 552 VLVVVGDPAAE-----RASARRILVPIIG-----LEYSFAAADL 586
```

Search completed: October 23, 2004, 03:58:40
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 03:57:47 ; Search time 85 Seconds
(without alignments)
1443.583 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGSKDRPPVLVS.....HLHYRLPADVRPPIISVRK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1364641 seqs, 323758627 residues

Total number of hits satisfying chosen parameters: 1364641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2005	100.0	379	14	US-10-007-527A-2
2	2005	100.0	379	14	US-10-007-452-2
3	2005	100.0	379	16	US-10-415-562A-2
4	609.5	30.4	459	14	US-10-007-527A-21
5	609.5	30.4	459	14	US-10-007-452-21
6	609.5	30.4	459	16	US-10-415-562A-21
7	325	16.3	427	9	US-09-835-381-6
8	322	16.1	451	14	US-10-007-527A-24
9	322	16.1	451	14	US-10-007-452-24
10	322	16.1	451	16	US-10-415-562A-24
11	318	15.9	427	9	US-09-835-381-8
12	318	15.9	427	14	US-10-196-232-7
13	316	15.8	456	14	US-10-007-527A-22

14	316	15.8	456	14	US-10-007-452-22	Sequence 22, Appl
15	316	15.8	456	16	US-10-415-562A-22	Sequence 22, Appl
16	282	14.1	466	9	US-09-826-191-9	Sequence 9, Appl
17	282	14.1	466	14	US-10-263-666-9	Sequence 9, Appl
18	282	14.1	466	14	US-10-261-481-9	Sequence 9, Appl
19	282	14.1	466	14	US-10-261-942-9	Sequence 9, Appl
20	266.5	13.3	528	14	US-10-007-527A-23	Sequence 23, Appl
21	266.5	13.3	528	14	US-10-007-452-23	Sequence 23, Appl
22	266.5	13.3	528	16	US-10-415-562A-23	Sequence 23, Appl
23	109.5	5.5	409	15	US-10-425-114-65990	Sequence 65990, A
24	103.5	5.2	417	14	US-10-156-761-10798	Sequence 10798, A
25	103	5.1	647	14	US-10-021-660-94	Sequence 94, Appl
26	102	5.1	907	16	US-10-437-963-156771	Sequence 156771, A
27	101	5.0	373	16	US-10-437-963-152208	Sequence 152208, A
28	101	5.0	823	16	US-10-437-963-156772	Sequence 156772, A
29	100.5	5.0	544	14	US-10-156-761-8540	Sequence 8540, Ap
30	99.5	5.0	468	14	US-10-369-493-12194	Sequence 12194, A
31	98.5	4.9	532	15	US-10-282-122A-43657	Sequence 43657, A
32	98.5	4.9	5835	15	US-10-378-083-20	Sequence 20, Appl
33	98	4.9	533	16	US-10-437-963-149861	Sequence 149861, A
34	97.5	4.9	515	16	US-10-437-963-187892	Sequence 187892, A
35	97	4.8	559	14	US-10-369-493-4607	Sequence 4607, Ap
36	97	4.8	567	14	US-10-369-493-7364	Sequence 7364, Ap
37	97	4.8	713	13	US-10-014-717-11	Sequence 11, Appl
38	96.5	4.8	938	16	US-10-437-963-189869	Sequence 189869, A
39	96	4.8	7746	14	US-10-156-761-7965	Sequence 7965, Ap
40	95.5	4.8	3798	13	US-10-014-717-6	Sequence 6, Appl
41	95.5	4.8	4471	14	US-10-205-032-10	Sequence 10, Appl
42	95	4.7	184	16	US-10-437-963-181278	Sequence 181278, A
43	95	4.7	321	16	US-10-437-963-183714	Sequence 183714, A
44	94.5	4.7	552	14	US-10-369-493-15887	Sequence 15887, A
45	94.5	4.7	552	16	US-10-437-963-135632	Sequence 135632, A

ALIGNMENTS

RESULT 1

US-10-007-527A-2
; Sequence 2, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CIL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rhodococcus AN12
US-10-007-527A-2

Query Match 100.0%; Score 2005; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSVSAEHLGSKDRPPVLVSDDKRGIRHRLPKLQIITSETNACGRPTSGVNGVTIYN 60

Db 1 MTSVSAEHLGSKDRPPVLVSDDKRGIRHRLPKLQIITSETNACGRPTSGVNGVTIYN 60

QY 61 GPKSGFGGLRSCGKWCIFCCAGKVGAGRADEISQVAFQIQTGSGVAMVTMTMRHTAGQ 120

Db 61 GPKSGFGGLRSCGKWCIFCCAGKVGAGRADEISQVAFQIQTGSGVAMVTMTMRHTAGQ 120

QY 121 RLHDLWTGLSAWAATNGRWRTEREMYGCDGYVRAVEITHGKNGWHVHALLMFSGD 180

Db 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVRRPPIISVRK 379
Db 361 LHVRLPAADVRRPPIISVRK 379
RESULT 2
US-10-007-452-2
; Sequence 2, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rhodococcus AN12
US-10-007-452-2
Query Match 100.0%; Score 2005; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
Db 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
QY 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
Db 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
QY 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
Db 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVRRPPIISVRK 379
Db 361 LHVRLPAADVRRPPIISVRK 379
RESULT 3
US-10-007-527A-21
; Sequence 21, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
US-10-007-527A-21
Query Match 100.0%; Score 2005; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
Db 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
QY 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
Db 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
QY 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
Db 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVRRPPIISVRK 379
Db 361 LHVRLPAADVRRPPIISVRK 379
RESULT 4
US-10-007-527A-21
; Sequence 21, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
US-10-007-527A-21
Query Match 100.0%; Score 2005; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
Db 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
QY 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
Db 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
QY 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
Db 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVRRPPIISVRK 379
Db 361 LHVRLPAADVRRPPIISVRK 379

Db 361 LHVRLPAADVRRPPIISVRK 379
RESULT 3
US-10-415-562A-2
; Sequence 2, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rhodococcus AN12
US-10-415-562A-2
Query Match 100.0%; Score 2005; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
Db 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
QY 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
Db 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
QY 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
Db 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVRRPPIISVRK 379
Db 361 LHVRLPAADVRRPPIISVRK 379
RESULT 4
US-10-007-527A-21
; Sequence 21, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
US-10-007-527A-21
Query Match 100.0%; Score 2005; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.9e-182;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
Db 1 MTSVSAEHLGKDRPPVLVSSDKRGIRHELPRKLOIITSETFNACGRPISGVNGVTIVN 60
QY 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
Db 61 GPKSGFGGLRSCGKGMICPCACGKVGARADEISQVVAHQLTGSGVAMVTMTMRHTAGQ 120
QY 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
Db 121 RLHDLWTGLSAAWKAATNGRRWTERMYGCDGVRAVEITHGKNGWHVHALLMFSGD 180
QY 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
Db 181 VSENLFSFSDAMFDRWTSKLVSLGFAAPLRNSGLDVRKIGGADQVLAAYLTKIAGV 240
QY 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
Db 241 GMEVSGDGKSGRGNRAPWEIADVGGDPQALELWREFEFGSMGRRRAIWSRGLRARA 300
QY 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
Db 301 GLGAELTDAQIVEGESAPVWVAIIPARSWMMIRTCAPIVFGELGLVEAGATWENLRDH 360
QY 361 LHVRLPAADVRRPPIISVRK 379
Db 361 LHVRLPAADVRRPPIISVRK 379

```
; SEQ ID NO 21
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Arcanobacterium pyogenes
US-10-007-527A-21

Query Match      30.4%; Score 609.5; DB 14; Length 459;
Best Local Similarity 36.3%; Pred. No. 4.5e-49;
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;

QY 5 SAHLSGKDRPPVLVSSD-----KRGIRHELKPKLQIITSETFNACGR-PISGVNGV 56
Db 67 SATHPLGNTVLTFFVSNESKKTAKRSRSEYELRDGLAEISTIESVRKCGRPVAPLVSL 126

QY 57 TIVNGPKSGFGILRSCGKWI CPCAGKVGARADEISQVVAHQLTG-SVAMVTMTMR 115
Db 127 RAKSDGKAGYGGIHTCGSVWACPVCSAKIAARRKTDLQQVVDHAVKHGTMVSMILTQR 186

QY 116 HTAGORLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL 175
Db 187 HHKQGLKHLWDALSTAWNRVTSGRRMTFEKQFGLGVYVRANEITHGKHGWHVSHVLI 246

QY 176 MFGSDVSENILESF-----SDAMPDRWTSKLVSLGFAAPLRNSGG 215
Db 247 ISEKD---PLTSTFVYQKQGRRLPYPPPEIYMSSDFAERWEAGLAKHGVDL-LRDSGG 302

QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVGMEVSGDGKSGRHNAPWEIAVDA 266
Db 303 LDWTAVKADARAIGN-----YVSKMOTSDAISSEVTLGFGFKARNGNRTFPQILADI 354

QY 267 VG-GDQOALELWRPEFGSMGRRRAIAWSRGLRABAGLAELTDAQIVQEESAPVMVAII 325
Db 355 LSLGDVDDLKWLKEYEKASFRRALTWKGLRDWANLGVQESDEEIA-SEEIGDEAIALF 413

QY 326 PARSMWMTIRTCAPYVFG--EILGLVEAGA-----TWNELRDHLHYRLP 366
Db 414 THDAWRQVRR-----FGAAELLDVTESGGRAAAYRWLDFRE-IDWSLP 455

RESULT 6
US-10-415-562A-21
; Sequence 21, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont De Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415.562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Arcanobacterium pyogenes
US-10-415-562A-21

Query Match      30.4%; Score 609.5; DB 16; Length 459;
Best Local Similarity 36.3%; Pred. No. 4.5e-49;
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;

QY 5 SAHLSGKDRPPVLVSSD-----KRGIRHELKPKLQIITSETFNACGR-PISGVNGV 56
Db 67 SATHPLGNTVLTFFVSNESKKTAKRSRSEYELRDGLAEISTIESVRKCGRPVAPLVSL 126

QY 57 TIVNGPKSGFGILRSCGKWI CPCAGKVGARADEISQVVAHQLTG-SVAMVTMTMR 115
Db 127 RAKSDGKAGYGGIHTCGSVWACPVCSAKIAARRKTDLQQVVDHAVKHGTMVSMILTQR 186

QY 116 HTAGORLHDLWTGLSAAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNGHVVHALL 175
Db 187 HHKQGLKHLWDALSTAWNRVTSGRRMTFEKQFGLGVYVRANEITHGKHGWHVSHVLI 246

QY 176 MFGSDVSENILESF-----SDAMPDRWTSKLVSLGFAAPLRNSGG 215
Db 247 ISEKD---PLTSTFVYQKQGRRLPYPPPEIYMSSDFAERWEAGLAKHGVDL-LRDSGG 302

QY 216 L-----DVRKIGGEADQVLAAYLTKI---ASGVGMEVSGDGKSGRHNAPWEIAVDA 266
Db 303 LDWTAVKADARAIGN-----YVSKMOTSDAISSEVTLGFGFKARNGNRTFPQILADI 354

QY 267 VG-GDQOALELWRPEFGSMGRRRAIAWSRGLRABAGLAELTDAQIVQEESAPVMVAII 325
Db 355 LSLGDVDDLKWLKEYEKASFRRALTWKGLRDWANLGVQESDEEIA-SEEIGDEAIALF 413

QY 326 PARSMWMTIRTCAPYVFG--EILGLVEAGA-----TWNELRDHLHYRLP 366
Db 414 THDAWRQVRR-----FGAAELLDVTESGGRAAAYRWLDFRE-IDWSLP 455

RESULT 5
US-10-007-452-21
; Sequence 21, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US NA
; CURRENT APPLICATION NUMBER: US/10/007.452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Arcanobacterium pyogenes
US-10-007-452-21

Query Match      30.4%; Score 609.5; DB 14; Length 459;
Best Local Similarity 36.3%; Pred. No. 4.5e-49;
Matches 148; Conservative 62; Mismatches 133; Indels 65; Gaps 15;

QY 5 SAHLSGKDRPPVLVSSD-----KRGIRHELKPKLQIITSETFNACGR-PISGVNGV 56
Db 67 SATHPLGNTVLTFFVSNESKKTAKRSRSEYELRDGLAEISTIESVRKCGRPVAPLVSL 126

QY 57 TIVNGPKSGFGILRSCGKWI CPCAGKVGARADEISQVVAHQLTG-SVAMVTMTMR 115
Db 127 RAKSDGKAGYGGIHTCGSVWACPVCSAKIAARRKTDLQQVVDHAVKHGTMVSMILTQR 186
```

```

; SEQ ID NO 24
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Streptomyces nigirifaciens
US-10-007-527A-24

Query Match      16.1%; Score 322; DB 14; Length 451;
Best Local Similarity 30.8%; Pred. No. 9.4e-22;
Matches 101; Conservative 41; Mismatches 110; Indels 76; Gaps 15;

Qy 49 PISGV-----NGVTIVNGPKSGFGGLRSCGKGMWICPCACGKAGVGAHRADEISQVVAHQ 102
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 PASGVIVAQTAAGTSVVL-----GLMRCGRILWCPVCAATIRHKRAEIBITAAVVEMI 54
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 103 GTGSVA-MVTMTMRHTAGQRLDHLMTGLSAWKAATNGRR-----WRTE----- 145
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 55 KRGGTAYLVTFTRHGHGTDRLADLMDALQGTRKTADAPRRPGAYQRLITGGTWAGRRAXD 114
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 146 -----REMYGCDGVYRAVEITHGK-NGWVHVHVALMFGSDV-----SENILESF 189
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 115 GHRAADREGIDRIGYGMIRATEVTVGQINGWPHIHAIVLVGRTGERSAKQIVGTF 174
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 190 --SDAMFDRWTSKLVSLGFAAPLR-----NSGGLDVRKIGIGHEAD-QVLAAYLFKIA 237
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 175 EPSEALDEWQGWRAV-WTAALRKVNPFPTDDRHGVDFKRLTERDANDLAEYIAKTQ 233
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 238 SG--VGMVEVSGDGKSGRHGNRAPEI---AVDAVGG---DPQA-----LELWREFE 281
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 234 DGKAPALELARADLUKTANGGNVAFPELLGRIGDLTGGMTEDDAAGVSGSLEWNLARWHEY 293
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 282 FSGMRRRAIAWSRGLRARAGLGAELTDA 309
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 294 RATKGRRAIEWTRYLRQMLGLDGGDTGA 321
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 9
US-10-007-452-24
; Sequence 24, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Streptomyces nigirifaciens
US-10-007-452-24

Query Match      16.1%; Score 322; DB 14; Length 451;
Best Local Similarity 30.8%; Pred. No. 9.4e-22;
Matches 101; Conservative 41; Mismatches 110; Indels 76; Gaps 15;

Qy 49 PISGV-----NGVTIVNGPKSGFGGLRSCGKGMWICPCACGKAGVGAHRADEISQVVAHQ 102
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 PASGVIVAQTAAGTSVVL-----GLMRCGRILWCPVCAATIRHKRAEIBITAAVVEMI 54
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 103 GTGSVA-MVTMTMRHTAGQRLDHLMTGLSAWKAATNGRR-----WRTE----- 145
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 55 KRGGTAYLVTFTRHGHGTDRLADLMDALQGTRKTADAPRRPGAYQRLITGGTWAGRRAXD 114
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 146 -----REMYGCDGVYRAVEITHGK-NGWVHVHVALMFGSDV-----SENILESF 189
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 115 GHRAADREGIDRIGYGMIRATEVTVGQINGWPHIHAIVLVGRTGERSAKQIVGTF 174
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 190 --SDAMFDRWTSKLVSLGFAAPLR-----NSGGLDVRKIGIGHEAD-QVLAAYLFKIA 237
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 175 EPSEALDEWQGWRAV-WTAALRKVNPFPTDDRHGVDFKRLTERDANDLAEYIAKTQ 233
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 238 SG--VGMVEVSGDGKSGRHGNRAPEI---AVDAVGG---DPQA-----LELWREFE 281
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 234 DGKAPALELARADLUKTANGGNVAFPELLGRIGDLTGGMTEDDAAGVSGSLEWNLARWHEY 293
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 282 FSGMRRRAIAWSRGLRARAGLGAELTDA 309
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 294 RATKGRRAIEWTRYLRQMLGLDGGDTGA 321
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:


```

```

US-09-835-381-6
; Sequence 6, Application US/09835381
; Patent No. US20020045223A1
; GENERAL INFORMATION:
; APPLICANT: SUGA, Mikiko
; APPLICANT: ASAKURA, Yoko
; APPLICANT: MORI, Yukiko
; APPLICANT: ITO, Hisao
; APPLICANT: KURAHASHI, Osamu
; TITLE OF INVENTION: ARGININE REPRESSOR DEFICIENT STRAIN OF CORYNEFORM BACTERIUM AND METHOD OF INVENTION: FOR PRODUCING L-ARGININE
; FILE REFERENCE: 206018USO
; CURRENT APPLICATION NUMBER: US/09/835,381
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: JP2000-129167
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Brevibacterium lactofermentum
US-09-835-381-6

Query Match      16.3%; Score 326; DB 9; Length 427;
Best Local Similarity 29.1%; Pred. No. 3.7e-22;
Matches 98; Conservative 58; Mismatches 131; Indels 50; Gaps 15;

Qy 34 LQQTITSTFNACGR-----PISGVNGVTIVNGPKSGFGGLRSCGKGMWICPCACGV 86
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 27 MYKITNSKALAGCHWRDRDEAVVSWSS-----NG--ASQFEGQLQNSHRWGSPLAELEV 79
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 87 GAHRADEISQVVAHQLTGSGVAVTM-TMRHTAGQRLDHLMTGLSAWKAATNGRRWRTE 145
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 80 MGERRIELAIAITKNHLAAGGALMFEVGVTRNRSOSFAQVEAGIKTAYSSWVKTSQWKE 139
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 146 REMYGCDGVYRAVEITHG-KNGWVHVHVALMFGSDVSENILESFSDAMFDRWTSKLVSL 204
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 RARYGVHTSYDYETDWSWANGHLHRNMLLFLDRPLSDDELKAFEDSMFSRWSAGVKA 199
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 205 GFAAPLRNSG-GLD-VRKIGGADQVLAAYLTKIASGVMEVSGDGSGRHGNRAPEI 262
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 200 GMDAPLREHGVKLDQVSTWGGDAK-MATYLAK---GMSQELTGSATKTASKGTYTPQM 255
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 263 AVDAVGGDPQALE-----LWREFEFGMERRAIAWSRGLRARAGLGAELTDAQI--- 311
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 256 -LDMLADQSDAGEDMDAULVAVRWREYEVGSKNLR-SWSRG--AKRALGIDYIDADVRE 311
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 312 VEQE-----ESAPYVVAIIPARSMWMT 335
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 312 MEELYKLAGLEAPERVESTRAVALVKPDDWKLIQS 348
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 8
US-10-007-527A-24
; Sequence 24, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97

```


Db 27 MYKITNSKALAGCHRRERDEAVSWSS-----NG--ASOFEGLQNSHRSWSSLAELV 79

QY 87 GAHRADEISQVVAHQLTGSAVMVTM-TMRHTAGQRLHDLWTGLSAWKAATNGRRWRTE 145

Db 80 MGERREIATATXNHLAAGALMMFVCTVHRNSQSFAQVEAGIKTAYSSWKTQWKKE 139

QY 146 REMYGCDDYVRAVEITHG-KNGWHVHVHALLMFSGDVSENLISFSDAMFDRWTSKLVS 204

Db 140 RARYGVHTYSDYEVTSWANGWHLHRNMLFLDRPLSDDELKAFEDSMFSRSAGVYKA 199

QY 205 GFAAPLRNSG-GLD-VPKIGCEADQVLAAYLTKTASGVGMEVSGDGKSGRHNAPWEI 262

Db 200 GMDAPLREHGKLDQVSTWGTGDAK-WATYLA-----GMSQELTGSATKTASKSYTFQM 255

QY 263 AVDAVGDDPOALE-----LWREFFBFGSMGRRAIAWSRGLRARAGLGAELTDAQI--- 311

Db 256 -LMDAQSDAGEDMDAVLARWEYEVGSKNLS-SWSRG--AKRALGIDYIDADVRE 311

QY 312 VEQE-----ESAPVVAIIPARSMWIRT 335

Db 312 MEELYKLAGLEAPERVESTRAVALVKPDDDKLIQS 348

RESULT 13

US-10-007-527A-22

Sequence 22, Application US/10007527A

Publication No. US20030044807A1

GENERAL INFORMATION:

APPLICANT: Tomb, Jean-Francois

APPLICANT: Bramucci, Michael G.

APPLICANT: Cheng, Qiong

APPLICANT: Kostichka, Kristy N.

TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

FILE REFERENCE: CL1709 US NA

CURRENT APPLICATION NUMBER: US/10/007,527A

CURRENT FILING DATE: 2001-12-05

PRIOR APPLICATION NUMBER: 60/254,868

PRIOR FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Microsoft Office 97

SEQ ID NO 22

LENGTH: 456

TYPE: PRT

ORGANISM: Streptomyces lividans

US-10-007-527A-22

Query Match 15.8%; Score 316; DB 14; Length 456;

Best Local Similarity 30.5%; Pred. No. 3.6e-21;

Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

QY 49 PISGV-----NGVTIVNGPKSGFGGLRSCGKWTCCPCAGKVGHAHRADEISQVVAHQ 102

Db 3 PASGVIVAQTAAGTSVVL-----GLMRCGRIMLCPVCAATIRHKRAEITAAVVEWI 54

QY 103 GTGSVA-MVTMTMRHTAGQRLHDLWTGLSAWKAATNGRR-----WRTE---- 145

Db 55 KRGGTAYLVTFTRARHGHTDLADLMDALQGTTRKTPDSPRRPGAYQRLITGGTWAGRAKD 114

QY 146 -----REMYGCDGVYRAVEITHGK-NGWHVHVHALLMFSGDV-----SENILESF 189

Db 115 GHRAADREGIRDRIGVGMIRATEVTVGQINGWHPHIIHLVVGRTGERSAKQIVATF 174

QY 190 --SDAMFDRWTSKLVSGLFAAPLR-----NSGGLDVRKIGGEAD-OVLAAYLTAKIA 237

Db 175 EPTGAALDEWQHRSV-WTAALRKVNPAFTPDDRHGVDVFKLETERDANDLAEYIAKTQ 233

QY 238 SG--VGMVEVSGDGKSGRHNAPWEI---AVDAVGG---DPOA-----LELWREPE 281

Db 234 DGKAPALELARADLKTATGNNVAPPELLGRIGDLTGGMTEDDAAGVGSLEWNLRSWHEYE 293

QY 282 FGSNGRRRAIAWSRGLRARAGLGAELTDA 309

Db 294 RATGRRRAIEWTRVLRQMLGLOGDTEA 321

RESULT 14

US-10-007-452-22

Sequence 22, Application US/10007452

Publication No. US20030093701A1

GENERAL INFORMATION:

APPLICANT: Tomb, Jean-Francois

APPLICANT: Bramucci, Michael G.

APPLICANT: Cheng, Qiong

APPLICANT: Kostichka, Kristy N.

TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

FILE REFERENCE: CL1709 US NA

CURRENT APPLICATION NUMBER: US/10/007,452

CURRENT FILING DATE: 2001-11-08

PRIOR APPLICATION NUMBER: 60/254,868

PRIOR FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Microsoft Office 97

SEQ ID NO 22

LENGTH: 456

TYPE: PRT

ORGANISM: Streptomyces lividans

US-10-007-452-22

Query Match 15.8%; Score 316; DB 14; Length 456;

Best Local Similarity 30.5%; Pred. No. 3.6e-21;

Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

QY 49 PISGV-----NGVTIVNGPKSGFGGLRSCGKWTCCPCAGKVGHAHRADEISQVVAHQ 102

Db 3 PASGVIVAQTAAGTSVVL-----GLMRCGRIMLCPVCAATIRHKRAEITAAVVEWI 54

QY 103 GTGSVA-MVTMTMRHTAGQRLHDLWTGLSAWKAATNGRR-----WRTE---- 145

Db 55 KRGGTAYLVTFTRARHGHTDLADLMDALQGTTRKTPDSPRRPGAYQRLITGGTWAGRAKD 114

QY 146 -----REMYGCDGVYRAVEITHGK-NGWHVHVHALLMFSGDV-----SENILESF 189

Db 115 GHRAADREGIRDRIGVGMIRATEVTVGQINGWHPHIIHLVVGRTGERSAKQIVATF 174

QY 190 --SDAMFDRWTSKLVSGLFAAPLR-----NSGGLDVRKIGGEAD-OVLAAYLTAKIA 237

Db 175 EPTGAALDEWQHRSV-WTAALRKVNPAFTPDDRHGVDVFKLETERDANDLAEYIAKTQ 233

QY 238 SG--VGMVEVSGDGKSGRHNAPWEI---AVDAVGG---DPOA-----LELWREPE 281

Db 234 DGKAPALELARADLKTATGNNVAPPELLGRIGDLTGGMTEDDAAGVGSLEWNLRSWHEYE 293

QY 282 FGSNGRRRAIAWSRGLRARAGLGAELTDA 309

Db 294 RATGRRRAIEWTRVLRQMLGLOGDTEA 321

RESULT 15

US-10-415-562A-22

Sequence 22, Application US/10415562A

Publication No. US20040115661A1

GENERAL INFORMATION:

APPLICANT: E.I. du Pont De Nemours and Company

TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

FILE REFERENCE: CL1709 US PCT

CURRENT APPLICATION NUMBER: US/10/415,562A

CURRENT FILING DATE: 2003-11-17

PRIOR APPLICATION NUMBER: 60/254,868

PRIOR FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Microsoft Office 97

SEQ ID NO 22

LENGTH: 456

TYPE: PRT

ORGANISM: Streptomyces lividans

US-10-415-562A-22

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 03:58:47 ; Search time 5309 Seconds

(without alignments)
3375.927 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005

Sequence: 1 MTSVSAEHLGSKDRPPVLVS.....HLHYRLPAADVRPPIISVRK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO.spool/US10007527/runat_22102004_114116_26617/app_query.fasta_1.519
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10007527@cgn_1_1_3731@runat_22102004_114116_26617 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2005	100.0	1140	6	AX548642
C 2	2005	100.0	6334	1	AY178757 Rhodococc
C 3	2005	100.0	6334	6	AX548646 Sequence
C 4	2005	100.0	9652	6	AX548648 Sequence

C 5	2005	100.0	11241	6	AX548647
C 6	2005	100.0	11241	12	AY180162
C 7	609.5	30.4	2439	1	APU83788
C 8	569	28.4	3540	1	AY150274
C 9	497	24.8	2051	1	PF66662
C 10	453	22.6	2661	1	SLMPG33A
C 11	452.5	22.6	9367	1	SN243257
C 12	425.5	21.2	4603	1	AY172684
C 13	409	20.4	300717	1	AE017229
C 14	399.5	19.9	2297	6	E17316
C 15	392.5	19.6	5750	1	AF085719
C 16	374.5	18.7	5804	7	AF428260
C 17	360.5	18.0	12855	1	AF533985
C 18	346.5	17.3	10992	1	STMPSN22
C 19	345	17.2	3748	12	AB032347
C 20	345	17.2	3784	12	AB032348
C 21	345	17.2	5128	1	AB084167
C 22	342.5	17.1	4109	1	CCA308231
C 23	337	16.8	5349	12	AY072037
C 24	337	16.8	6342	12	AY072040
C 25	337	16.8	6607	12	AY072039
C 26	337	16.8	7560	6	AR475430
C 27	337	16.8	8058	12	AY072038
C 28	337	16.8	8215	12	PI414669
C 29	337	16.8	8830	1	PIJ101CG
C 30	337	16.8	12738	6	AR475429
C 31	333.5	16.6	2297	1	AF273219
C 32	326	16.3	3147	12	AF092035
C 33	326	16.3	4447	1	AF092037
C 34	326	16.3	4447	6	BD106940
C 35	326	16.3	4447	6	BD106940
C 36	326	16.3	4447	6	BD139585
C 37	326	16.3	4457	1	BLPLASML
C 38	326	16.3	5135	12	SVE133193
C 39	326	16.3	5216	12	SVE133194
C 40	326	16.3	6241	12	AF092036
C 41	326	16.3	6272	12	AF091268
C 42	326	16.3	6398	12	AF092931
C 43	326	16.3	6601	12	SVE133195
C 44	326	16.3	8161	12	AY585307
C 45	318	15.9	4447	6	AR173115

ALIGNMENTS

RESULT 1	AX548642	Sequence 1 from Patent WO02055709.	1140 bp	DNA	linear	PAT 27-NOV-2002
AX548642	Sequence 1 from Patent WO02055709.					
LOCUS	AX548642					
DEFINITION	AX548642					
ACCESSION	AX548642					
VERSION	AX548642.1	GI:25813612				
KEYWORDS						
SOURCE	Rhodococcus erythropolis					
ORGANISM	Rhodococcus erythropolis					
REFERENCE	1					
AUTHORS	Bramucci,M.G., Cheng,Q., Kostichka,K.N. and Tomb,J.F.					
TITLE	Rhodococcus cloning and expression vectors					
JOURNAL	Patent: WO 02055709-A 1 18-JUL-2002;					
FEATURES	E.I. DU PONT DE NEMOURS AND COMPANY					
source	Location/Qualifiers					
	1. .1140					
	/organism="Rhodococcus erythropolis"					
	/mol_type="unassigned DNA"					
	/strain="ANI2"					
	/db_xref="taxon:1833"					

ORIGIN

Alignment Scores:
Pred. No.: 6 77e-123 Length: 1140
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 100.00% Indels: 0			
DB: 6 Gaps: 0			
US-10-007-527A-2 (1-379) x AX548642 (1-1140)			
Qy	1	MetThrSerValSerAlaGluHisLeuSerGlyValAspArgProProValLeuValSer	20
Db	1	ATGACCAAGCGTAAGTGTGTAACACCTTTCCGGCAAGACCGGCTCCCGCTCGTGTGC	60
Qy	21	SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer	40
Db	61	TCCGATAGCGCGCATCCGCGACGAACCTCGACCCCAACTTCACAAATCACCACTCA	120
Qy	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60
Db	121	GAACAATTTAACGCTGTGGCGCGCATTTTCGGCTGAACGGTGTGACCAATTTGTCAAC	180
Qy	61	GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysPro	80
Db	181	GGTCCGAAGAGTTCTGGATTTCGGAGGCTTCGTTCCTGCGGAAAGGCTGGATCGCCC	240
Qy	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValValAlaHis	100
Db	241	TGCTGTGCGGAAAGTCGGTGCACATCGTCGACAGCAATTTCTCAAGTTGTGCTCAT	300
Qy	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	301	CAACTCGGACTGGATCTGTTCGATGTGACGATGACCATCGCGCCATACAGCTGGTCAG	360
Qy	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg	140
Db	361	CGGCTCACGACCTATGACTTGGACTTTCGGCAGCCTGGAAGCTGCAACACGGTCGT	420
Qy	141	ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160
Db	421	CGTTGGCTAGCAAGCTGAATGTACGGCTCGCAGCATACGTGCGGCTGTTGAAATC	480
Qy	161	ThrHisGlyLysAsnGlyTyrPheHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	481	ACTCACGAAAAAAGCGGTGGCGTCCACGCTTCCGCGCTTACTCATGTTCAGTGGTGAC	540
Qy	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys	200
Db	541	GTGAGTGAGAACATCTCGAATCTTTCGGATGGATGTTTCGATCGGTGACTTCCAA	600
Qy	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220
Db	601	CTCGTATCTCTGGGATTTGCTGGCCACTACGTAATTCGGGTGCTCGATGTACGAAAG	660
Qy	221	IleGlyGlyGluAlaAspGlnValLeuAlaIleValLeuThrLysIleAlaSerGlyVal	240
Db	661	ATCGCGCGTGAAGCTGATCAAGTTCCTCGCTCGTATCTGACGAAATTCATCTGGCGTT	720
Qy	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr	260
Db	721	GGTATGGAGGTTGGTAGTGGCGACGAAAGAGTGTGCATGGCAACCGTGACCCCTGG	780
Qy	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe	280
Db	781	GAATTCGCTGTGTGATGCATGGCGGGGATCCACAAAGCTTGGAACTGTGGCGAGAAAT	840
Qy	281	GluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla	300
Db	841	GAGTTTGGTTTCATGGGACGTGGGCAATTCGGTGGTCCCGGGATTCGGTCCCGAGCT	900
Qy	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal	320
Db	901	GGTCTTGGCGGAGAACTACAGATGCTCAGATCGTTGTAGCAGGAATCTGCCCGGTC	960
Qy	321	MetValAlaIleIleProAlaArgSerTyrMetMetIleArgThrCysAlaProTyrVal	340
Db	961	ATGGTTGCGATCATTTCCGCGCGCATCGTGGATGATGATTCGGACTTGTTCGCGCTTACGTC	1020

Qy	341	PheGlyGluIleLeuGlyLeuValGluAlaGlyValaThrTyrGluAsnLeuArgAspHis	360
Db	1021	TTGCGCGAGATCTCGACTCGTCGAAGCTGGCGGACCTTGGGAAATCTCTCGTGATCAC	1080
Qy	361	LeuHisTyrArgLeuProAlaAlaAspValArgProIleIleSerValaArgLys	379
Db	1081	TTGCATATCGATTGCCCGCAGCGATGTGCGCCCCCGATAATATATCGGTTCCGCAAG	1137
RESULT 2			
AY178757/c			
LOCUS			
Rhodococcus erythropolis cryptic plasmid pAN12, complete sequence.			
ACCESSION			
AY178757			
VERSION			
AY178757.1 GI:28628256			
KEYWORDS			
Rhodococcus erythropolis			
SOURCE			
ORGANISM			
Rhodococcus erythropolis			
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
Corynebacterineae; Nocardiaceae; Rhodococcus.			
REFERENCE			
AUTHORS			
Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and			
Cheng, Q.			
TITLE			
A small cryptic plasmid from Rhodococcus erythropolis:			
characterization and utility for gene expression			
Appl. Microbiol. Biotechnol. 62 (1), 61-68 (2003)			
MEDLINE			
22718480			
PUBMED			
12835922			
REFERENCE			
2 (bases 1 to 6334)			
AUTHORS			
Kostichka, K., Tao, L., Bramucci, M., Tomb, J.-F., Nagarajan, V. and			
Cheng, Q.			
TITLE			
Direct Submission			
JOURNAL			
Submitted (11-NOV-2002) CE&D, E.I. Dupont de Nemours Inc.,			
Experimental Station, Wilmington, DE 19880-0328, USA			
FEATURES			
source			
1. 6334			
/organism="Rhodococcus erythropolis"			
/mol_type="genomic DNA"			
/db_xref="taxon:1833"			
/plasmid="pAN12"			
/note="cryptic plasmid"			
complement (1912..3051)			
/gene="rep"			
complement (1912..3051)			
/gene="rep"			
/note="Rep"			
/codon_start=1			
/transl_table=11			
/product="replication protein"			
/protein_id="AAO31602.1"			
/db_xref="GI:28628257"			
/translation="MTSVAEHLGSKDRPPVLVSSDKRGIRHRLPKLQQTSETFN			
ACGRPIGVNGVTIVNGPVGKSGFGSLSCGKGMCPCCAGKVGARHDEISQVVAHOL			
GTGSVAVMTMTMRHTAGRLHDLMWGLSAWKAATNGRCRRMTERMGDCGIVRAVEI			
THGKNGHVVHALLMFSGDVSENIKFSDFMDFRMTSKLVSLGFAAPLRNSGGLDV			
RKIGCEADQVLAAYLTIKIAGVGMEVSGSGKSGRGNRAPEWIAVAVGDPQALEL			
WRPEFGSMGRRAIWSRGLRARAGLGAELTDAQIVOEESAPVAVVAIPARSWMIR			
TCAPYVGEILGLVEAGATWENLRDLHLHYLRPAADVRRPIISVRK"			
complement (4288..5178)			
/gene="div"			
complement (4288..5178)			
/gene="div"			
/note="Div"			
/codon_start=1			
/transl_table=11			
/product="putative cell division protein"			
/protein_id="AAO31603.1"			
/db_xref="GI:28628258"			
/translation="WDQTDTPITAIQWNEIAPQVMDIAKDAHWLIQKTRSGKSCQ			
TYNLAQAGSNPAVRVVGVDPTSVLLAPFVRRPAEPNIEIQLNDPDKVLVLFQVKA			
ESDRIECFWDRRIDKISLSPALPLIILVLEEPGIIIEGADQDFATNGLKPADRYAP			
RITSLVRQIAQSAKAGIRMLLQAQRASISVGGNARSNFVAKVTLRVDEPESVKMLH			
PNATPECALVEGEVFGGFFDQPLGRKQMIRTVKVGSEYISTYASTVSNADLAYEALN			

IDRAQRMTIASYPHLDIG"1

ORIGIN

Alignment Scores:

Pred. No.: 5,35e-122 Length: 6334

Score: 2005.00 Matches: 379

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 1 Gaps: 0

US-10-007-527A-2 (1-379) x AV178757 (1-6334)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20

Db 3051 ATGACCGGTAAGTCTGAACACCTTCCGGCAAGACCGGCCCTCCCGTCTCGTGTCG 2992

QY 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40

Db 2991 TCCGATAGCGCGCATCCGGCAGCACTGGACCCCAAACTTCAACAAATCACACGTC 2932

QY 41 GluThrPheAsnAlaCysGlyArgProLysLeuSerGlyValAsnGlyValThrIleValAsn 60

Db 2931 GAAACATTTACGCTGTGGCGCGCGGATTTCTGGCGTGAACGGGTGACCATTTGTC 2872

QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80

Db 2871 GGTCGGAAGGTTCTGGATTCGGAGGCTTCGTTCTCGGAAAGGGCTGGATTCGCC 2812

QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValAlaHis 100

Db 2811 TGCTGTGGGAAAGTCGGTGCACATCGTCAGCAAGAAATTTCTCAAGTTGTGTCAT 2752

QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetArgHisThrAlaGlyGln 120

Db 2751 CAACTCGGACTGGATCTGTTGCCATGGTACGATGACCATGCCCATACAGCTGGTCAG 2692

QY 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140

Db 2691 CGGCTCCACGACCTATGAGCTGGACTTCGGCAGCTTGGAAAGCTGGACCAACGGTCGT 2632

QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160

Db 2631 CGTTGGCGTACGGAACGTGAATCTCTCGATCGCATGCGTACGCGTGTGTAATC 2572

QY 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180

Db 2571 ACTACGGAAGAAACGGCTGGCAGCTCCAGCTTCCAGCGCTACTCATGTTCACTGGT 2512

QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200

Db 2511 GTGAGTGAGAACATCTCTCGAATCTTCTCGATCGCATGTTGATCGGTGATCTCCAA 2452

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220

Db 2451 CTCGTATCTCTGGGATTTGCTGGCGCACTACGTAATTCGGTGGTCTCGATGACGAA 2392

QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240

Db 2391 ATCGCGGTGAAGCTGATCAAGTTCTCGCTGGCTATCTGACGAAATTTGATCTGGCG 2332

QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260

Db 2331 GGTATGGAGTTGGTAGTGGCAGCGAAGAAAGTGGTCGATCGCAACCGTGACCCCTGG 2272

QY 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280

Db 2271 GAAATCGCTGTTGATGCGTGGCGGGATCCACAGCGTTGAACTGTGGCGAGATTT 2212

QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAla 300

Db 2211 GAGTTTGGTTGATGGAGCTCGGGCAATCGCGTGTGTCGATTCGCTGGCGCGAGCT 2152

QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320

Db 2151 GGTCCTTGGCGCAGAACTAACGATCTCAGATCGTTGAGCAGGAAGATCTCCCGCGTC 2092

QY 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal 340

Db 2091 ATGTTTGGATCATTTCCGGCGGATCGTGGGATGATTCGGACTGTGGCGCTTACGTC 2032

QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360

Db 2031 TTCGGCGAGATCTCGGACTCGTCAAGCTGGCGGACTTGGAAAAATCTTCGTGATCAC 1972

QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379

Db 1971 TTGCATTATGATTCGCCGCGAGATGTGGCGCCCGGATAATATCGTTGCGCAAG 1915

RESULT 3

AX548646/c

LOCUS AX548646 6334 bp DNA linear PAT 27-NOV-2002

DEFINITION Sequence 5 from Patent WO205709.

ACCESSION AX548646

VERSION AX548646.1 GI:25813614

KEYWORDS Rhodococcus erythropolis

SOURCE Rhodococcus erythropolis

ORGANISM Rhodococcus erythropolis

REFERENCE 1

AUTHORS Bramucci, M.G., Cheng, Q., Kostichka, K.N. and Tomb, J.F.

TITLE Rhodococcus cloning and expression vectors

JOURNAL Patent: WO 0205709-A 5 18-JUL-2002;

E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES

source

1. .6334

/organism="Rhodococcus erythropolis"

/mol_type="unassigned DNA"

/strain="AN12"

/db_xref="taxon:1833"

ORIGIN

Alignment Scores:

Pred. No.: 5,35e-122 Length: 6334

Score: 2005.00 Matches: 379

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-007-527A-2 (1-379) x AX548646 (1-6334)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20

Db 3051 ATGACCGGTAAGTCTGAACACCTTCCGGCAAGACCGGCCCTCCCGTCTCGTGTCG 2992

QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40

Db 2991 TCCGATAGCGCGCATCCGGCAGCACTGGACCAACCTTCAACAAATCACACGTC 2932

QY 41 GluThrPheAsnAlaCysGlyArgProLysLeuSerGlyValAsnGlyValThrIleValAsn 60

Db 2931 GAAACATTTACGCTGTGGCGCGCGATTTCTGGCGTGAACGGGTGACCATTTGTC 2872

QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80

Db 2871 GGTCGGAAGGTTCTGGATTCGGAGGCTTCGTTCTCGGAAAGGGCTGGATTCGCC 2812

QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValAlaHis 100

Db 2811 TGCTGTGGGAAAGTCGGTGCACATCGTCAGCAAGAAATTTCTCAAGTTGTGTCAT 2752

QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetArgHisThrAlaGlyGln 120

Db 2751 CAACTCGGACTGGATCTGTTGCCATGGTACGATGACCATGCCCATACAGCTGGTCAG 2692

QY 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140

Db 2691 CGGCTCCACGACCTATGAGCTGGACTTCGGCAGCTTGGAAAGCTGGACCAACGGTCGT 2632

QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160

Db 2631 CGTTGGCGTACGGAACGTGAATCTCTCGATCGCATGCGTACGCGTGTGTAATC 2572

QY 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180

Db 2571 ACTACGGAAGAAACGGCTGGCAGCTCCAGCTTCCAGCGCTACTCATGTTCACTGGT 2512

QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200

Db 2511 GTGAGTGAGAACATCTCTCGAATCTTCTCGATCGCATGTTGATCGGTGATCTCCAA 2452

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220

Db 2451 CTCGTATCTCTGGGATTTGCTGGCGCACTACGTAATTCGGTGGTCTCGATGACGAA 2392

QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240

Db 2391 ATCGCGGTGAAGCTGATCAAGTTCTCGCTGGCTATCTGACGAAATTTGATCTGGCG 2332

QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260

Db 2331 GGTATGGAGTTGGTAGTGGCAGCGAAGAAAGTGGTCGATCGCAACCGTGACCCCTGG 2272

QY 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280

Db 2271 GAAATCGCTGTTGATGCGTGGCGGGATCCACAGCGTTGAACTGTGGCGAGATTT 2212

QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAla 300

Db 2211 GAGTTTGGTTGATGGAGCTCGGGCAATCGCGTGTGTCGATTCGCTGGCGCGAGCT 2152


```

Db      4152  GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGAGAAATCGCCCGGTC 4093
Qy      321  MetValAlaIleIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db      4092  ATGGTTGGGATCATTCGCGCGCGATCGTGGGATGATGATTCGGACTTGTGGCGCTTACGTC 4033
Qy      341  PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db      4032  TTCGGGAGATCCTCGGACTCGTGAAGCTGGCGGCACTTGGGAAATTCCTGIGATCAC 3973
Qy      361  LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
Db      3972  TTGCATTATCATGTTGCCGCGAGCATGTGGCGCCCGCATATATCGTTGGCAAG 3916

RESULT 5
AX548647/c
LOCUS      AX548647      11241 bp      DNA      linear      PAT 27-NOV-2002
DEFINITION Sequence 6 from Patent WO02055709.
ACCESSION  AX548647
VERSION     AX548647.1  GI:25813615
KEYWORDS    Shuttle vector pRHR17
SOURCE      Shuttle vector pRHR17
ORGANISM    Shuttle vector pRHR17
            artificial sequences; vectors.
REFERENCE   1
AUTHORS     Bramucci,M.G., Cheng,Q., Kostichka,K.N. and Tomb,J.F.
TITLE       Rhodococcus cloning and expression vectors
JOURNAL     Patent: WO 02055709-A 6 18-JUL-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES    Location/Qualifiers
             1..11241
             /organism="Shuttle vector pRHR17"
             /mol_type="unassigned DNA"
             /db_xref="taxon:214938"

ORIGIN
Alignment Scores:
Pred. No.:      1.07e-121      Length:      11241
Score:          2005.00      Matches:      379
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              6              Gaps:        0

US-10-007-527A-2 (1-379) x AX548647 (1-11241)
Qy      1  MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
Db      6641  ATGACCAGCGTAAGTGTGAACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTCG 6582
Qy      21  SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
Db      6581  TCCGATAAGCCGGCATCCGGCAGCAACTCGACCCCAAACTCAACAAATCACACGTC 6522
Qy      41  GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn 60
Db      6521  GAAACATTTAACGCTGTGGCCGCGCATTTCTGGCGTGACCGTGTGACCATGTCAAC 6462
Qy      61  GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db      6461  GGTCCGAAAGGTTCTGATTCGAGGCGCTTCGTTCTCGCGAAAGGCGTGAATCGCCC 6402
Qy      81  CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValValAlaHis 100
Db      6401  TGTGTGCGGAAAGTCGGTGACATCGTGCAGCAAAATTTCTCAAGTTGTGTGTCAT 6342
Qy      101  GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db      6341  CAACTCGGACTGGATCTGTGGATGTGTGACCATGACCATGCGCCATACAGCTGTGTCAG 6282
Qy      121  ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140

```

```

Db      6281  CGGCTCCACGACCTATGCACTGGACTTTCGGCAGCGCTGGAAGCTGCGACCAACGGTCGT 6222
Qy      141  ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db      6221  CGTTGGCGTACGGAACGTGAAATGTACGGCTGCGACGATACGTCGGCGCTGTGGAATC 6162
Qy      161  ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db      6161  ACTCAGGAAACACGGCTGCGACGTCCACGTTTCACGCGCTACTCATCTTCAGTGGTGAC 6102
Qy      181  ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db      6101  GTGAGTGAGACATCTCTCGAATCTTCTCGAATCGGATGTTTCGATCGGTGACATTCCTCAA 6042
Qy      201  LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db      6041  CTCGTATCTCTGGGATTTGCTCGGCCACTAGTAATTCGGGTGCTCTCGATGTACGAAAG 5982
Qy      221  IleGlyGlyGluAlaAspGlnValLeuAlaIleTyrLeuThrLysIleAlaSerGlyVal 240
Db      5981  ATCGCGCGTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTCATCTGGCGTT 5922
Qy      241  GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db      5921  GGTATGAGGTTGTTAGTGGCGCGAAAGTGTGACATGGCAACCGTCACCCCTGG 5862
Qy      261  GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db      5861  GAAATCGCTGTTGATCAGTGGCGCGGATCCACAGCGTTGGAACCTGCGCGAGAATTT 5802
Qy      281  GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAla 300
Db      5801  GAGTTGCTTTCGATGGGACCTCGGGCAATCGCGTGGTCCGCTGATTTGCGTCCGAGCT 5742
Qy      301  GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluGluSerAlaProVal 320
Db      5741  GGTCTTGGGCGAGAACTTACAGATGCTCAGATCGTTGAGCAGGAAGATTCGCCCGGTC 5682
Qy      321  MetValAlaIleIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db      5681  ATGGTTGCGATCATTCGCGCGCGATCGTGGATGATGATTCGACTTGTGGCGCTTACGTC 5622
Qy      341  PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db      5621  TTCGGCGAGATCTCGGACTCGTGAAGCTGGCGGCGACTTGGGAAATCTTCGTGATCAC 5562
Qy      361  LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
Db      5561  TTGCATTATCGATTGCCGCGAGCGGATGTCGGCCCCCGGATAATATATCGTTCCGCAAG 5505

RESULT 6
AX180162/c
LOCUS      AX180162      11241 bp      DNA      circular      SYN 23-JUL-2003
DEFINITION Shuttle vector pRHR17, complete sequence.
ACCESSION  AX180162
VERSION     AX180162.1  GI:30313707
KEYWORDS    Shuttle vector pRHR17
SOURCE      Shuttle vector pRHR17
ORGANISM    Shuttle vector pRHR17
            artificial sequences; vectors.
REFERENCE   1 (bases 1 to 11241)
AUTHORS     Kostichka,K., Tao,L., Bramucci,M., Tomb,J.-F., Nagarajan,V. and
            Cheng,Q.
TITLE       A small cryptic plasmid from Rhodococcus erythropolis:
            characterization and utility for gene expression
JOURNAL     Appl. Microbiol. Biotechnol. 62 (1), 61-68 (2003)
MEDLINE     22718480
PUBMED      12835922
2 (bases 1 to 11241)
AUTHORS     Kostichka,K., Tao,L., Bramucci,M., Tomb,J.-F., Nagarajan,V. and
            Cheng,Q.
TITLE       Direct Submission
JOURNAL     Submitted (14-NOV-2002) CR0D, E. I. Dupont de Nemours Inc.,

```

Experimental Station, Wilmington, DE 19880-0328, USA

FEATURES

Location/Qualifiers

1..11241
 /organism="Shuttle vector pRHBR17"
 /mol_type="genomic DNA"
 /db_xref="taxon:214938"
 /locus
 /note="Rhodococcus - Escherichia coli shuttle vector; Sep I linearize pAN12 cloned into the PvuII site of pBR328"
 join(9925..11241,1..3590)
 /organism="Cloning vector pBR328"
 /mol_type="genomic DNA"
 /db_xref="taxon:221946"
 3591..9924
 /organism="Rhodococcus erythropolis"
 /mol_type="genomic DNA"
 /db_xref="taxon:1833"
 /plasmid="pAN12"
 /note="cryptic plasmid"

ORIGIN

Alignment Scores:
 Pred. No.: 1.07e-121 Length: 11241
 Score: 2005.00 Matches: 379
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-10-007-527a-2 (1-379) x AY180162 (1-11241)

Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
 Db 6641 ATGACCGAGCTGAAGTCTGAACACCTTTCCGGCAAGACCGGCTCCGCTCGTGTGCG 6582
 Qy 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
 Db 6581 TCCGATAGCCGGCATCCGACGACGACTGGACCCCAACTTCAACAAATCACCACGTCA 6522
 Qy 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
 Db 6521 GAAACATTTAAGCGCTGTGGCGGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 6462
 Qy 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrIleCysPro 80
 Db 6461 GGTCCGAAAGGTTCTGGATTCTGGAGCCTTCTCTCGGAAAGGGCTGGATCTGCCCC 6402
 Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
 Db 6401 TGCTGTGGGAAAGTCTGGTGACATCTGCAGACGAAATTTCTCAAGTTGTGCTCAT 6342
 Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
 Db 6341 CAACCTCGGAGCTGGATCTGTTGGATGTGACGATGACCATGGCGCATACAGCTGGTCAG 6282
 Qy 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaIleTyrLysAlaAlaThrAsnGlyArg 140
 Db 6281 CGGTCCACGACCTATGACGTGGATTTCTGGCAGCTTGGAAAGCTGGCAACACGGTCTGT 6222
 Qy 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 Db 6221 CGTTGGCGTAGCGAAGCTGAATGTACGGCTGGCGAGCATGCTGGCGGCTGTGAATC 6162
 Qy 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 Db 6161 ACTCAGCAAAAACCGGTGGCACCTCCACGTTCCAGCGCTACTCATGTTCAGTGGTGCAC 6102
 Qy 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
 Db 6101 GTGAGTGAACATCTCTGAATCTCTCGGATGCGATGTTCTGATCGGTGGACTTCCAAA 6042
 Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLeuAspValArgLys 220

Db 6041 CTCGTATCTCTGGGATTTGCTGGCCACTACGTAATTCGGGTGCTCTCGATGACGAAG 5982
 Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
 Db 5981 ATCGCGCGTGAAGCTGATCAAGTTCTCGCTGCCATCTGACGAAATTTGATCTGGCGTT 5922
 Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
 Db 5921 GGTATGAGGTGGTAGTGGCGACGGAAGTGGTCGACATGCAACCGTCGACCCCTGG 5862
 Qy 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
 Db 5861 GAAATCGCTGTTGATGCAAGTGGCGGGATCCCAAGCGTTGGAACCTGTGGCGAGAATTT 5802
 Qy 281 GluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla 300
 Db 5801 GAGTTTGGTTTCGATGGGACGTCGGCAATCGCGTGGTCCCGTGAATTCGCGCCGAGCT 5742
 Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 Db 5741 GGTCTTGGGCGAGACTTACAGATGCTCAGATCGTTGAGCAGGAAGATCTGCCCGGTC 5682
 Qy 321 MetValAlaIleIleProAlaArgSerTyrMetMetIleArgThrCysAlaProTyrVal 340
 Db 5681 ATGTTTCGATCATTCGCGCGCATCGTGGATGATGATTCGGACTTGTGGCCTTACGTC 5622
 Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTyrGluAsnLeuArgAspHis 360
 Db 5621 TTGCGCGAGATCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 5562
 Qy 361 LeuHisTyrArgLeuProAlaAlaAspValArgProIleIleSerValArgLys 379
 Db 5561 TTGCATTATCGATTGCCCGCAGCGGATGTGCGCCCCCGATAATATCGTTCCGCAAG 5505
 RESULT 7
 APU83788 2439 bp DNA circular BCT 15-JUN-1998
 LOCUS Actinomycetes pyogenes plasmid pAP1, complete plasmid sequence.
 DEFINITION U83788
 ACCESSION U83788
 VERSION U83788.1 GI:1805288
 KEYWORDS Arcanobacterium pyogenes
 SOURCE Arcanobacterium pyogenes
 ORGANISM Arcanobacterium pyogenes
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Actinomycineae; Actinomycetaceae; Arcanobacterium.
 REFERENCE 1 (bases 1 to 2439)
 AUTHORS Billington, S.J., Jost, B.H. and Songer, J.G.
 TITLE The Arcanobacterium (Actinomycetes) pyogenes plasmid pAP1 is a member of the pJ101/pJVI family of rolling circle replication plasmids
 JOURNAL J. Bacteriol. 180 (12), 3233-3236 (1998)
 MEDLINE 98292760
 PUBMED 9620977
 REFERENCE 2 (bases 1 to 2439)
 AUTHORS Billington, S.J., Jost, B.H. and Songer, J.G.
 TITLE Direct Submission
 JOURNAL Submitted (06-JAN-1997) Veterinary Science, University of Arizona, 1117 East Lowell Street, Tucson, AZ 85721, USA
 FEATURES
 source
 1..2439
 /organism="Arcanobacterium pyogenes"
 /mol_type="genomic DNA"
 /strain="BBR1"
 /db_xref="taxon:1661"
 /plasmid="pAP1"
 rep_origin 1..110
 /note="single stranded origin; putative"
 rep_origin 39..44
 /note="site of single strand initiation; putative"
 gene 221..1600
 /gene="rep"
 CDS 221..1600
 /gene="rep"
 /function="replication protein"

```

/codon_start=1
/transl_table=11
/product="Rep"
/protein_id="AAC46399.1"
/db_xref="GI:1805289"
/translation="MNRLSRRTALSLPARQIQKIPVPAAGGRSLKSPFGMTATWSARGG
ASDSDRRKRSQIPSNRRGRSATHPLGNTVLTFPVNSKSKTKTAKRSRVERVELRDG
LAEITTESVKRGKRVPAVLVSLRAKSKDGKAGYGLHTCGSWACGVCSAKIAARR
KTLQOQVHVHAKVGMTVSLTLTQRHHKQGLKHLWDALSTANWRVTSGRRWIEFKR
QFGLGVVVRANEITHGKHGHWVHSHVLIISSEKDLSTFVYVYKQGRRLRFPYPEIYM
SSDFIARWGAAGHGVDFLRDGGDLDTVAKDARAIGNYVSKMOTSDAISSEVTL
GGFKARNGRNTFPQILADILSLGDVLDLKEVYKASFGRRALTWSKGLRDWANLG
VEGSDERIASIEEIGDEAIALFTHDARQVRFRFGAAELLDVTESGRRAAAYRWLDFRI
DWSLPPKIE"
1673. 1676
RBS
CDS
1685. .2023
/note="ORF112"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC46400.1"
/db_xref="GI:1805290"
/translation="MFFIGKRTKFOPIAPCKPSIPDGNGKRTGFSSEKMRTOAIT
EDRLVELTSIPQFVQLIDDAVKSRELLFEKVNLAVALPRNGGISTYLTGEAVHV
PAPVVEVSR"
2010. .2013
RBS
CDS
2020. .2409
/note="ORF129"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC46401.1"
/db_xref="GI:1805291"
/translation="MMSEPYGISNABELQLVFRPAKREARLWMSRLKKELESKHS
TPTEALESCEKLYWVFEPRLVDVILHEVGECRCAGTNSAVARSLSRAFIAPFVQS
LENGGRNPAQLTQDQVAGMRKTLGR"
2378. .2406
/note="double stranded origin; putative"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 7.38e-31 Length: 2439
Score: 609.50 Matches: 148
Percent Similarity: 51.47% Conservative: 62
Best Local Similarity: 36.27% Mismatches: 133
Query Match: 30.40% Indels: 65
DB: 1 Gaps: 15

```

US-10-007-527A-2 (1-379) x APU83788 (1-2439)

```

QY 5 SerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSerSerAsp----- 22
Db 419 TCAGGACCATCCCTTGGCAATACGGTATTAACTTTTCCAGTATCAACAGAGCAAG 478
QY 23 -----LysArgGlylleArgHisGluLeuArgProLysLeuGlnGlnle 37
Db 479 AAAACAGCTAAATCTCGCGTCTTGAGAGATACGAACACTCAGACCGATTAGCCGAAATC 538
QY 38 ThrThrSerGluThrPheAsnAlaCysGlyArg---ProIleSerGlyValAsnGlyVal 56
Db 539 TCGACCATTGAGTCGCGCGGAAGTGTGGCCGGTGGCCGTCGTCACCTCTCGTCTCGTTG 598
QY 57 ThrIleValAsnGlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGly 76
Db 599 CGAGCAAAATCTGACGGTAAGCGCGGATATGGTGTTCACACTTGTGGAAGCGTC 658
QY 77 TrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGln 96
Db 659 TGGCGTGCCTCAGTCTGTAGCGGAAATTCGCCGCTCGCCGAAACACGACCTCAACAG 718
QY 97 ValValAlaHisGlnLeuGlyThrGly---SerValAlaMetValThrMetMetArg 115
Db 719 GTCGTTGACACCGCCGTAAACACAGGAATGACCGTCTCAATGCTTACGCTACCCAGCGT 778

```

```

QY 116 HisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTrpLysAla 135
Db 779 CACCACAAGGACAAAGGCTAAACACCTCTGGGACGCTTGTGACGGCATGGAATCGC 838
QY 136 AlaThrAsnGlyArgAspGlyTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrVal 155
Db 839 GTTACCTCTCTGCTGCTGTGGATTGAGTTCAAGAGCAAAATTGTGTTTAGTCGGTTATGTT 898
QY 156 ArgAlaValGluIleThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeu 175
Db 899 CGAGCCAATGAATTACTCATGAAACACGCGTGGCATGTCATTCCTCATGTTCTGATT 958
QY 176 MetPheSerGlyAspValSerGluAsnIleLeuGluSerPhe----- 189
Db 959 ATTTCCGAGAAAGAC-----CCGCTGACCAGCACGTTTGTCTATCAACGCAACAA 1009
QY 190 -----SerAspAlaMetPheAsp 195
Db 1010 GAGCGCGCGCCTTCCCTACCCCCCCAGAGATTATATGTCATCCGATTTTCATTGCTGAA 1069
QY 196 ArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGly 215
Db 1070 CGGTGGGAAGCTCGCCTTGGGAAGCAGCGGCTGATT---CTCCCGAATTCGGAGGC 1126
QY 216 Leu-----AspValArgLysIleGlyGlyGluAlaAspGlnValLeu 229
Db 1127 TTGGACTGGACCGCTTGCAGAAAGACGCGCGAGCCATCGGCAAC----- 1168
QY 230 AlaAlaTyrLeuThrLysIle-----AlaSerGlyValGlyMetGluValGlySer 246
Db 1169 -----TATGTCAGCAAGATGTCAGACGCTCCACAGCGGATTAGTCGGAAGTCACGTTG 1222
QY 247 GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAla 266
Db 1223 GCGCGCTTCAAAAACCCGAAACGGAACAGACGCGCTTCAGATATCTCGCGGATATC 1282
QY 267 ValGly---GlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPheGlySerMet 285
Db 1283 CTTTCGCTCGCGCATGTCGACGACCTCAAGCTCTGAAAGAAATATGAGAAAGCTTCGTTT 1342
QY 286 GlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGlu 305
Db 1343 GGAACGCGTGCACCTTACATGCTCGAAAGGCGCTCAGAGATTGGGCAAAATCTCGCGTTGAA 1402
QY 306 LeuThrAspAlaGlnIleValGluGlnGluSerAlaProValMetValAlaIleIle 325
Db 1403 CAGTCGACGAGAGAGATTGCC---TCTGAGAAATCGGGGACGAGCAATAGCGTATTT 1459
QY 326 ProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrValPheGly-----Glu 343
Db 1460 AGCATGACGCTTGGCGTCAGGTGCGACGT-----TTTGGAGCGCGTGAA 1504
QY 344 IleLeuGlyLeuValGluAlaGlyAla-----ThrTrpGluAsnLeuArg 358
Db 1505 CTACTCGATGTCACGAATCCGAGGTCGTGCGCGCGCTTACCGCTGTTGGATTTTAGG 1564
QY 359 AspHisLeuHisTyrArgLeuPro 366
Db 1565 GAA---ATTGATTGTCATTGCGCT 1585

```

RESULT 8

```

AY150274 3540 bp DNA circular BCT 02-JAN-2003
LOCUS Propionibacterium granulosum cryptic plasmid pPG01, complete
DEFINITION sequence.
ACCESSION AY150274
VERSION AY150274.1 GI:27465054
KEYWORDS Propionibacterium granulosum
SOURCE Propionibacterium granulosum
ORGANISM Propionibacterium granulosum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Propionibacterineae; Propionibacteriaceae; Propionibacterium.

```

1 (bases 1 to 3540)
Farrar,M.D. and Holland,K.T.
Isolation and characterisation of a cryptic plasmid from the human
skin commensal *Propionibacterium granulosum*
Unpublished

2 (bases 1 to 3540)
Farrar,M.D. and Holland,K.T.
Direct Submission
Submitted (13-SEP-2002) Skin Research Centre, Division of
Microbiology, University of Leeds, Leeds LS2 9JT, UK
Location/Qualifiers

FEATURES
source
1. 3540
/organism="Propionibacterium granulosum"
/mol_type="genomic DNA"
/strain="PF283"
/db_xref="taxon:33011"
/plasmid="cryptic plasmid pPG01"
74..280
/codon_start=1
/transl_table=11
/product="Orf3"
/protein_id="AA078121.1"
/db_xref="GI:27465055"
/translation="MGQGYGLNRGFYNGSHNEIIGGYPMGRVIWVWPDPAII.SAMPLE
SSGFSFGHVPEAEVFRHCDHD"
449..1792
/gene="tra"
449..1792
/gene="tra"
449..1792
/gene="tra"
/note="plasmid transfer protein"
/codon_start=1
/transl_table=11
/product="Tra"
/protein_id="AA078122.1"
/db_xref="GI:27465056"
/translation="WAGERTVLTGVWMLRILPVLVLLGVAFACWCFSHLVVYVIGFA
VAALAAEPFPPVQVQKMQVPARYVYRWMSLAKAFKPIESGSDRIYYRPGLOMI
RSDRHVHLVLPVAPGLADSAAYLEKGAEEIQRLRGKKSWRTCIKVPABHGLDIIPR
DATAGDELPAFSPVNPVGVKPGSEVVDLSHPSHILVSGKTSRGKSSFFVYGLL
DQMRHLPTVAVGDPPTGILNFDLGDGWDGDLRSKRITNDADAARVQVLSMTIDEMD
RRIYLNCBHRDKWNRNDFSPGRRLLIIVILEYPLIERLQNFDSARGARSDDRA
SKAAGLIVAYEGAKVGVLLVTPREDAKIIGPLPRAQTTTRVTFQAQSDGLRSH
PELSSQVQKNSWASGVGEIADGVIPLTRFRSRAELTDLHRPGASVVGIDILIQ"
1979..3280
/gene="rep"
1979..3280
/gene="rep"
/note="plasmid replication protein"
/codon_start=1
/transl_table=11
/product="Rep"
/protein_id="AA078123.1"
/db_xref="GI:27465057"
/translation="MLNIVTILSKAOPHDLCSPEPCARSERQRAAGDRLGTTRAT
FPVDSPCPKRNEGRHRYEMRDLGRNPFQWELERVKRGCAVPVSQRIALMAGHGAG
YAGLATCGSWACIPVCAAKISAHRRDELARVVQVAVGLGVKVSQRIALMAGHGAG
ELWASLQSGNVAIVTSGRWQEFCAQLGVQWYKAVEVTHGSHGWHVHVILVSKQDP
TSVDTKIRRRQKGRRTPEVQRPEDFIAERWSRGKRGVDFIAGSGGLDQWTA
DSGDEALGRYAKVMNSVDGLANRATLGGFKARRGNRTPFQILEDPLDTGSETDLR
LWPTYVYASGKRALTWSGLRDMAGMESEMSDEOVAAQDQGEAVALFDHDAWRQIR
TAGAALLLELELHGSEGYAWLKKRRHYEPLVPWSTSI"

Alignment Scores:
Pred. No.: 5,28e-28 Length: 3540
Score: 569.00 Matches: 141
Percent Similarity: 51.78% Conservative: 63
Best Local Similarity: 35.79% Mismatches: 12
Query Watch: 39 Indels: 39
DB: 1 Gaps: 14

US-10-007-527A-2 (1-379) x AY150274 (1-3540)

```

Db      3155 ACTGCGGTCGCGCTTCTCTCGACGAGCTGGAGCTCCACGGATCCGAGGCGCTCTAC 3214
QY      353 ThrTgGluAsnLeuArgAspHisLeuHisTyrArgLeuPro 366
Db      3215 GCCTGGCTGAGAGACGA---AGATCCATTATGAGATACCT 3253

RESULT 9
LOCUS   PFR6662/c
DEFINITION Propionibacterium freudenreichii plasmid pLME108 rep gene.
ACCESSION AJ006662
VERSION  AJ006662.1 GI:3212127
KEYWORDS putative; rep gene.
SOURCE  Propionibacterium freudenreichii
ORGANISM Propionibacterium freudenreichii
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.

REFERENCE
AUTHORS Dasen,G.H., Miescher,S., Teuber,M. and Meile,L.
TITLE Molecular analysis of plasmid pLME108, a plasmid isolated from
Propionibacterium freudenreichii
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2051)
TITLE Direct Submission
JOURNAL Submitted (03-JUN-1998) Dasen G.H., Institute of Food Science,
Laboratory of Food Microbiology, ETH Zurich, LFO G24.2,
Schmelzbergstr. 9, CH-8092 Zurich, SWITZERLAND
FEATURES
source location/Qualifiers
1. 2051
/organism="Propionibacterium freudenreichii"
/mol_type="genomic DNA"
/db_xref="taxon:1744"
/lab_host="DF2"
/plasmid="pLME108"
/complement (223. 1434)
/gene="rep"
/complement (223. 1434)
/gene="rep"
/codon_start=1
/transl_table=11
/product="putative rep protein"
/protein_id="CAA07175.1"
/db_xref="GI:3212128"
/db_xref="GOA:O69764"
/db_xref="TrEMBL:O69764"
/translation="MVVDARSRGVRGCGGSPAVAPLGGALDNHANNVSPAPAA
GEVADKRHRFSVRVWLMWEHTSLKRVAFCGRVAASAVASVGRVCSDDGRAGPAGLQSC
GSVWPCVNAKIMARRGELGAAVETWTKHGRVAFMTFTVHRSKDSLTAVWDGVA
SGWRRVTSKGWTSDDLRRHGVGVFVVEVTHGRNGWHVHLVFLVGVDFGDLALH
RSMFGWGRGVLAAGTGPLARAQDYQOMSAATGLDLHARYLSKAOFQKIGHELNTS
QSKTARSALSTRWEVLGDAANGLAKEVLMFEWKEGRGRQRQIGWSAGLRDLRLGM
VRESDDVIAAREVGSVADTVALITGDGWRRLVGQQLYECLRSCELGGQAGLARWLQE
HGIEHLEVEVD"

ORIGIN
Alignment Scores:
Score: 1,46e-23 Length: 2051
PDB: 497.00 Matches: 127
Percent Similarity: 50.80% Conservative: 63
Best Local Similarity: 33.96% Mismatches: 144
Query Match: 24.79% Indels: 40
DB: 1 Gaps: 15

US-10-007-527A-2 (1-379) x PFR6662 (1-2051)

QY      19 ValSerSerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThr 38
Db      1296 GTGGCTGCTGACAGGCGCAACACCGCTTTTCGGTCCGTTATGTTGGCGGTATACG 1237
QY      39 ThrSerGluThrPheAsnAlaCysGlyArg---ProIleSerGlyValAsnGlyValThr 57

```

```

Db      1236 TCGTTGAAGCGGTTGCGTTCTGCGGTCTGTGTTCGCGCTCGCGGTTCGCTTCGTTGGT 1177
QY      58 IleValAsnGlyProLysGlyPheGlyLeuArgSerCysGlyLysGlyTyr 77
Db      1176 GTGCGGTGCAGCGACGCGCGCGGTTTCGTTGGTTGCAGAGTTCGGATCGGTCTGG 1117
QY      78 IleCysProCysCysAlaGlyValGlyAlaHisArgAlaAspGluIleSerGlnVal 97
Db      1116 GCCTGCCCGGTGTGTACGCAAGATCATGCTCGCGCTGGCTTGAACCTCGCGCGCG 1057
QY      98 Val-----AlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThr 113
Db      1056 GTCGAGACCTCGACTAAGCAC-----GGTGTCTGTGTGGCTTCATACGTTACG 1006
QY      114 MetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAlaIleTyr 133
Db      1005 GTTCGGCACTCTCGTAAGGATTCCTTGACGCGCTGTGGGATGTGTGCTTCGGGGTGG 946
QY      134 LysAlaIleThrAsnGlyArgArgTyrThrGluArgGluMetTyrGlyCysAspGly 153
Db      945 CGCGAGTCACTTCGGGCAAGGGCTGCACGCTCCACAGCTCCGTCACGGGTTCGAGGCG 886
QY      154 TyrValArgAlaValGluIleThrHisGlyLysAsnGlyTyrHisValHisAla 173
Db      885 TTCGTGCTGTGTGTCGAGGTTCACGCGGGCGCAATGGTGGCATGTGCATCTCATGTG 826
QY      174 LeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMet 193
Db      825 CTGTGTCTCTGTCGCGACTTCGGGACGCTCTG-----GGCTGCACCGGTCCATG 772
QY      194 PheAspArgTyrThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSer 213
Db      771 TTCGCGCGTTGGGAACGCGCGCTGCTCGCTGCTGGCTCGGACTCCCTTCGTCGCG 712
QY      214 GlyGlyLeuAspValArgLysIleGlyGlyGluAlaAsp---GlnValLeuAlaIleTyr 232
Db      711 -----CAAGACGTTACGACATGAGCGGACAACTGGCTGTGATCACCTTCGCGCGCTAC 658
QY      233 LeuThrLys-----IleAlaSerGlyValGlyMetGluValGlySerGlyAspGlyLys 250
Db      657 CTGAGCAAGCGCGCAGTTCAGGGGAAGATCGGCACAGTTCAGGTTGAGAACTCCAGAGCAAG 598
QY      251 SerGlyArgHis-----GlyAsnArgAlaProTyrGluIleAlaValAspAlaValGly 268
Db      597 ACGGCGCGTTCTGCGCTCTCGACTCGACTTGGAGGTCTGGGCGCGCGGCCAAC 538
QY      269 GlyAspProGlnAlaLeuGluLeuTyrArgGluPheGluPheGlySerMetGlyArgArg 288
Db      537 GGTTCGCGCAAGAGGTTCGTTCTCTGTTGAGTGGGAGAAGGTTCAGCGGGACGCGAGA 478
QY      289 AlaIleAlaTyrSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAsp 308
Db      477 CAGATTGGCTGTTCTCGGGGTTGAGGACCGCTCTCGGCTCATGTTGAGGAATCTGAC 418
QY      309 AlaGlnIleValGluGlnGlu---GluSerAlaProValMetValAlaIleIleProAla 327
Db      417 GACGTGATCGCGCTGAGGAAGTTCGGTCTGTGTCACACATGCTGCTTTCATCACGGGT 358
QY      328 ArgSerTyrMet-----MetIleArgThrCysAla 337
Db      357 GACGATGCGCTGCGCTGTTGGGGCAGCAGAGCTCTACGAGTGTCTCGAAGTTGC--- 301
QY      338 ProTyrValPheGlyGluIleLeuGlyLeuValGluAlaGly---AlaThrTyrGluAsn 356
Db      300 -----GAATTGGCGGT-----CAAGCTGGCTTCGCGCGGTG----- 268
QY      357 LeuArgAspHis---LeuHisTyrArgLeuProAlaAlaAsp 369
Db      267 CTGCAAGACGACGGAATCGAACACCAATGTTGGAGGTGGAC 226

RESULT 10
SLMPG33A
LOCUS

```

S.lavendulae miniplasmid pSLG33 sequence.
 X69872
 VERSION X69872.1 GI:47179
 KEYWORDS miniplasmid
 SOURCE Streptomyces lavendulae
 ORGANISM Streptomyces lavendulae
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Streptomycineae; Streptomycetaceae; Streptomyces.
 REFERENCE
 1 Pelsberg, J., Petricek, M. and Tichy, P.
 Nucleotide sequence of the mini-plasmid pSLG33 from Streptomyces
 lavendulae-grasseri RIA746
 JOURNAL Nucleic Acids Res. 21 (15), 3582 (1993)
 MEDLINE 93348001
 PUBMED 8346038
 REMARK (sites)
 REFERENCE
 2 (bases 1 to 2661)
 Pelsberg, J.
 Direct Submission
 AUTHORS Submitted (28-DEC-1992) J. Pelsberg, Inst. of Microbiology, Cell
 TITLE and Molecular Microbiology Div., Videnka 1083, Prague 4,
 JOURNAL CZECHOSLOVAKIA
 FEATURES
 source
 1..2661
 Location/Qualifiers
 /organism="Streptomyces lavendulae"
 /mol_type="genomic DNA"
 /strain="RIA746"
 /db_xref="taxon:1914"
 /dev_stage="mycelia"
 /plasmid="miniplasmid pSLG33"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,55e-20 Length: 2661
 Score: 453.00 Matches: 129
 Percent Similarity: 48.33% Conservative: 59
 Best Local Similarity: 33.16% Mismatches: 160
 Query Match: 22.59% Indels: 41
 DB: 1 Gaps: 16
 US-10-007-527A-2 (1-379) x SLMPG33A (1-2661)
 Qy 3 SerValSerAlaGluHisLeuSerGlyLysAspArgPro-----ProValLeu 18
 Db :::::|||||
 586 AATGTTTCCACCCCTCCCTCTCTTCCACTCAATCTCCAGCTCAGAGCGGATTCGCC 645
 Qy 19 ValSerSerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnLeuThr 38
 Db |||||:::|||||
 646 GTTCGTAGGGCGCGGGAACGGCGCTTCGCCAGCGGCGAGTCTCTGGACGATCACC 705
 Qy 39 ThrSerGluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGly----- 55
 Db |||||:::|||||
 706 AGCGTTCCTCCCGGTCGCAAGTCGCTTCGCGACCAATGGAGGCGTCTGTCAGCGCGCTG 765
 Qy 56 ---ValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGly 74
 Db :::::|||||
 766 GCGGCGATCGGTGTCACAGGC---TCGTCGCTACTGTCGCGCTTCGTCGCGCTGTCGCTG 822
 Qy 75 LysGlyTTPileCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIle 94
 Db |||||:::|||||
 823 TCCATCTGGGCTGCCCGTGTCTGTCGCGCAAGATCAGGCGCCACCGCGCGACGAGATC 882
 Qy 95 SerGlnValValAlaHisGlnLeuGlyThrGlySerValAla---MetValThrMetThr 113
 Db :::::|||||
 883 GCACGAGCGCGCGCAAGCACCTCGCGCGAGCGCGCGCGCTGGATGTCACGTGAGCG 942
 Qy 114 MetArgHisAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaTrp 133
 Db |||||:::|||||
 943 GCGCGACACAAGACACACAGCTTGAGCCCTCTTCGACCGCGTCCCGACGCGCTGG 1002
 Qy 134 LysAlaAlaThrAsnGlyArgArgTrpArg-----ThrGlu 145
 Db :::::|||||
 1003 CGCAGCTCTGTCCGGCGGGGCTTGGCGCGGAGACCCCAAGCGGGAGTCTCTCGCGGAA 1062

Qy 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGly---Lys 164
 Db |||||:::|||||
 1063 CGCGACCGGCTGGAGTCCGGGCAACATCCGCTCACTCGAAGTCACTTACGGAACCTCGT 1122
 Qy 165 AsnGlyTTPHisValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsn 184
 Db |||||:::|||||
 1123 AACGGITGGACCCGACCTTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1182
 Qy 185 IleLeuGluSerPheSerAspAlaMetPheAspArg---TrpThrSerLysLeuValSer 203
 Db :::::|||||
 1183 CTC-----GCCTACGCAATGCGCATGGGACAAACCTGGAGGGCTCGCTCAAGAAA 1236
 Qy 204 LeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGly 223
 Db |||||:::|||||
 1237 GCCGATTGCAACCGAGCAGGAACAC-----GGCATCACCTGGTCCAAAGTC----- 1284
 Qy 224 GluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGly---ValGlyMet 242
 Db :::::|||||
 1285 ACACCCCAAGAGAGCGAGCGAATACATCCCAAGGCTCAAGAGGAGGAGGTCGGAAC 1344
 Qy 243 GluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGlu--- 261
 Db |||||:::|||||
 1345 GAAATCGCCGAGGCGCATGAAGCGCGGTCTCGAACTCTGGCCCCGCTTCCGAAATG 1404
 Qy 262 IleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGlu 281
 Db |||||:::|||||
 1405 CTCGAATACTTCGGGACAGCGGTGACATGGCAGTGTCCCGATTTGGCAGGATGACGAG 1464
 Qy 282 PheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArg---Ala 300
 Db |||||:::|||||
 1465 AAGGGCACGTTCCCGCTCGCGCGATCACCTGGAGCGGTGCGCTTCGTCGCGAACTCCTC 1524
 Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 Db |||||:::|||||
 1525 GGACGAGAAAGAGAGTCCACGCAAGAACTTCAGCTGGAGAGATCGGAGGCGGAGCG 1584
 Qy 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
 Db |||||:::|||||
 1585 TGG---GCACCTCTCGCGCGCGAATCGCTCCGAGCGATCCGCCGT---GTCCCCGGTCTC 1638
 Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrp----- 354
 Db |||||:::|||||
 1639 CAGTCCCGCATCTCTCGACACCGCGAGAAACGCGGCTTCGTGGAACCTGTCACCTGCTGA 1698
 Qy 355 -----GluAsnLeuArgAsp 359
 Db :::::|||||
 1699 CTGCCTACCATCTTCTCTACGAGAT 1723
 RESULT 11
 SNA243257 9367 bp DNA circular BCT 02-DEC-1999
 LOCUS Streptomyces natalensis plasmid pSNA1 including spda, kora, spdb2
 DEFINITION and rep genes.
 ACCESSION AJ243257
 VERSION AJ243257.1 GI:6523476
 KEYWORDS KORA protein; rep gene; replication protein; SpdA protein; SPDB2
 protein; transfer protein.
 SOURCE Streptomyces natalensis
 ORGANISM Streptomyces natalensis
 Bacteria; Actinobacteria; Actinomycetales;
 Streptomycineae; Streptomycetaceae; Streptomyces.
 REFERENCE
 1 Mendes, M.V., Aparicio, J.F. and Martin, J.F.
 Complete nucleotide sequence and characterization of pSNA1 from the
 Pimaricin-producing Streptomyces natalensis that replicates by a
 rolling circle mechanism
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 9367)
 AUTHORS Mendes M.V.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUN-1999) Mendes M.V., Microbiology, Inbiotec, AV.
 REAL 1, Leon, 24006, SPAIN

FEATURES

```

source
    Location/Qualifiers
    1..9367
        /organism="Streptomyces natalensis"
        /mol_type="genomic DNA"
        /db_xref="taxon:68242"
        /plasmid="psN1"
        /country="South Africa"
        266..814
            /gene="spda"
        266..814
            /gene="spda"
        /function="putative intramycelial transfer protein"
        /note="ORF4"
        /codon_start=1
        /transl_table=11
        /product="SPDA protein"
        /protein_id="CAB62258.1"
        /db_xref="GI:6523477"
        /db_xref="GOA:Q9RIQ3"
        /db_xref="TrEMBL:Q9RIQ3"
        /translation="MYVTAPHTGSMPPSAGTVTGMTOGETPRDADGRRICEHCLVXP
        VPESLGNKPRSYCNRQRAYEARROGRFVKQTVNALLRRMMANATQATSRDNP
        EAKSRDAGQATSRDVASAPATRPAPAPRPPVQPTQAALEFEEDQADBEAP
        RPPARRMTDLXQMKLPUGFG"
        complement(1203..1940)
        /gene="kora"
        complement(1203..1940)
        /gene="kora"
        /function="putative transcriptional regulator"
        /note="ORF1"
        /codon_start=1
        /transl_table=11
        /product="KORA protein"
        /protein_id="CAB62259.1"
        /db_xref="GI:6523478"
        /db_xref="GOA:Q9RIQ2"
        /db_xref="TrEMBL:Q9RIQ2"
        /translation="MPKXERSAPPYMQVWTHFRTRIMSGELSPGDALPTDRGLAEFG
        VSRATQALTVLTKLELBAVOGSTRVQSAPLHNGHNDRAVSRTRRIYAGBY
        ARITTSLEPAPADBALGIEGAPAIRIRVTYNASDVALSMTSLWDGALSSTAP
        ALVSTRIRGTRWYLETKTRGVVQGDVRVRLATDEBAQLLSLEQPAAVKLTRTI
        LRDEGTLLEYGVSVTAGRESIYEYTV"
        2096..3025
            /gene="spdb2"
        2096..3025
            /gene="spdb2"
        /function="putative intramycelial transfer protein"
        /note="ORF2"
        /codon_start=1
        /transl_table=11
        /product="SPDB2 protein"
        /protein_id="CAB62260.1"
        /db_xref="GI:6523479"
        /db_xref="GOA:Q9RIQ1"
        /db_xref="TrEMBL:Q9RIQ1"
        /translation="NEVVRPHRPAGADPAQLERILRLGLTVLVVVAVGLAFTATNV
        TQFALDHSVPTGWTAMLDPLVASALGTVLVVDGLLSEFVRPGMASLLRWFGLATW
        LNMWCSLWPTGPGIPGIPHLDPAGLLHSVLPLVLAIAITTTTTRIRIVSKIAVID
        EADTCWTPVDTMPAPASAPVRSEPLTASAPAAPAPVRPQRELEVICGABQVFR
        PALPAPVPSPKPEPAPVPNGADDPHVAIVRAWMWMDGELSTROTAEVRVGLSKTTVA
        NIWRELAHSHGQRPAPGQPLPGAGQTVLQGVAA"
        6656..8086
            /gene="rep"
        6656..8086
            /gene="rep"
        /note="ORF3"
        /codon_start=1
        /transl_table=11
        /product="replication protein"
        /protein_id="CAB62261.1"
        /db_xref="GI:6523480"
        /db_xref="GOA:Q9RIQ0"
        /db_xref="TrEMBL:Q9RIQ0"

```

```

/translation="MTTDKALSRCRYALGGGVTPKVTGGTAYLAGLATCGKVHICP
CCGAKIRSARTVELQAGTAWEDIGINGPKIRPARAAWKDINDRTAEADWAAIBEAD
GLOAGGLGMLTLTMRHYSRHTLAEVLVTQORDAWKKALQONAGRDWRKAKDYGVVGFV
RAWEVTVGEANGWHPHWHVLFVFDKPLTPEOGDALEFVIYEAWSTALQDVGAFLPDRE
HGVRDLDSHGEGGPRAYLMKYQDKAAWTAAEMTRDTKAGRDGHRTPETIARVL
LTEDAADDDRAQTVRLMQEYETRAARMRLVNSGLRKLALVBLDTRTDEITAAEE
RQGEALAVILADPWHQHTARKRSLQLLKAEEKGGQDKVALVESWGLVMDRDLVPP
PVDLEAAQRAEELDRRGVGEADQRKAHRVAKVSAPPAARKASAAAKTKAPKAA
AHCTECGALDRLAGASGAGRHC"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 7,63e-20 Length: 9367
Score: 452.50 Matches: 128
Percent Similarity: 42.82% Conservative: 54
Best Local Similarity: 30.12% Mismatches: 158
Query Match: 22.57% Indels: 85
DB: 1 Gaps: 14

US-10-007-527A-2 (1-379) x SNA243257 (1-9367)
QY 19 ValSerSerAspLysArgGlyLeuArgProLysLeuGlnGlnThr 38
DB 6602 GTGCGGAGCATCGCGGATGGCGTCCGTCACGACTCCGTAGGCGCTGTACGACGTGACG 6661
QY 39 ThrSerGluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIle 58
DB 6662 ACGACCAAGCGCTCAGCGCGTGGCGCGGTACGCGCTCGT---GGCGGCTCACCCCC 6718
QY 59 ValAsnGlyProLysGlySerGlyPhe---GlyGlyLeuArgSerCysGlyLysGlyTrp 77
DB 6719 AAGGTGACCCCGCGCGAAGCGCTACCTCGCGGCTGGCCACCTGCGGCAAGGTGCAC 6778
QY 78 IleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnVal 97
DB 6779 ATCTGCCCTGCTCGCGCGGAGATCGGTCGCCCTGCTAGCTCGAATCCAGCGGCGCC 6838
QY 97 ----- 97
DB 6839 GGGACGCGCTGGGAGGACATCGCGGCGCATCAACGGCCGCAAGATCCGCCCGCGCGGCC 6898
QY 98 -----ValAla 99
DB 6899 GCCTGGAAGAGGACATCAACAGCCGCGGAGGACTGGGCGGCGCATCGAGAGGCC 6958
QY 100 HisGlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGly 119
DB 6959 GATGGGCTCCAGCGTGTGGCTCGGCTCGGCTGCTGACGCTGACCATGCGGCACTACTCCCG 7018
QY 120 GlnArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsn--- 138
DB 7019 CACACGCTCGCGAGCTGTGTACCCAGCGGCGACGCGGCGGCTTGGAAAGGCGCTCGGCCAG 7078
QY 139 -----GlyArgArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArg 156
DB 7079 AACCGGCGCGCGACTGGCGCAAGCGAAGAGGACTAGCGCGCTTGTGCGCTTCTCGTCCGG 7138
QY 157 AlaValGluIleThrHisGlyLys---AsnGlyTyrHisValHisValHisAlaLeuLeu 175
DB 7139 GCCTGGAGGTCACTACCGCGAAGCAACGCTGCGCACCGCGCTGCGGCTGCTGCTGCTGCT 7198
QY 176 MetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAsp 195
DB 7199 TTCTTCACACGCGCTGACCCCGGAGCGAGCGCGCTCGGAGAGGTTCATCTACGAG 7258
QY 196 ArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGly--- 214
DB 7259 GCATGGTCCACCGCGCTTCAGGACGCTGCGGCGCTACTCTCGCGCGCGTGAACACGCGCTC 7318
QY 215 GlyLeuAspValArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThr 234
DB 7319 CGCCTCGACCTGTCGGGCGCATGTTAGGCGGACCGCGG-----GCCGCGTACCTCATG 7372

```

QY	235	LysAlaAlaSerGly-----ValGlyMetGluValGlySerGlyAspGly	249	/db xref="GI:27657762"
Db	7373	AACTACGACGCGAAAGCGGCTGGACACGCGCGCGAGATGACGCGACGACG	7432	/translation="MSKOKLSPPRAKVIAAVDREVPVQVAHKIGSTROKAGSALMAL HXNLIDRFRCIYAPLEVDV" complement (613. .2061)
QY	250	LysSerGlyArgHisGlyAsnArgAlaProTrrpGluIleAla-----Val	264	/gene="repa" complement (613. .2061)
Db	7433	AAAGCTGGCGCGACGCGCCACCGCCGCTTACTGTCTCAACGCGCTACGCAAGCGCTCGCGCGCCCTG	7492	/gene="repa" /note="replication protein; essential for pag3 replication" /codon_start=1 /transl_table=11 /product="Repa" /protein_id="AAO18194.1" /db xref="GI:27657763"
QY	265	AspAlaValGlyGlyAsp---ProGlnAlaLeuGluLeuTrrpArgGluPheGluPheGly	283	/translation="MSILASKQAEHATSPNQATPATHTTGRIGTTDTTKHKVKALSS PALFRALMDLQRTMKVYEDTKLACGGRVAGGSVASLWQAGQPARWGGLQNSHS WMSGVASSVITCRRANEVDAAIQAWANGAGLTPAHYRGVSTSPSBRKQKQGVNVPV ERGISMTLTIRNSQSLTEVWDATAGCQWAVTNTAARGGATGADKRGYGLAHWY RAIEVTHGKGMHVHLHVLFDHRLVLSVDERDSLADRFDRWAAKAVRLGRAPSRDR GIDVHVAASSDDAKSTGGYCKGMLSGIAAETTTGGVTKKAGKDNRTPFOILGDLGK QYTKRDHALWEKSGKRRGTGWSQGTGKVLGINSDDQIDDSLGDNDDQSEVVAM VGRGWCALASDTKEAVIDLDAVAATDSAGAKAARDVLKLFVGAHTVTAIGLAKTD LDTVLPHDTDFKRELHEAVWSGRKQTALPVL" 2396. .2785
Db	7493	GATGCCCGGACGACGACCGTGGCCACCGTGGCGGTGTGGCAGGAGTACGACGCGG	7552	/note="orf3" /evidence=not experimental /transl_table=11 /product="hypothetical protein" /protein_id="AAO18195.1" /db xref="GI:27657764"
QY	284	SerMetGlyArgArgAlaIleAlaTrrpSerArgGlyLeuArgAlaArgAlaGlyLeuGly	303	/translation="MYRSATCLALHALNVVRCFGVSVSRETLDMTDYATLNRIKCMAT KAELOOLKDMQALKEQOOLAFNAREELEQORNLKPVVTVRVPQYVRDEMRLAAI CGWTLQFPAANAFAGMIDELSEQDQEH" complement (2840. .3232)
Db	7553	GCGCGCGGATGCGCGCGCTGTACTGTCTCAACGCGCTACGCAAGCGCTCGCGCGCCCTG	7612	/note="orf4" /codon_start=1 /evidence=not experimental /transl_table=11 /product="hypothetical protein" /protein_id="AAO18197.1" /db xref="GI:27657765"
QY	304	AlaGluLeu-----ThrAspAlaGluIleValGluGlnGluSerAlaProVal	320	/translation="MTKPVNRFPSANDIEAMIGLIVIANWENYCTDQKAKTAKAIAIA LNKWADERLRGKDLAHQDETHIERVLRHLWLFQKATHDNWQLDDEVHSLQSVHR IEAAGEFTRYGHNVGVVOLLAEFDTE" complement (3344. .4030)
Db	7613	GTCGAATCGACACCGCCACCGCGGAGATGCCCTGAGGAAACCGCCAGGCGGCC	7672	/codon_start=1 /evidence=not experimental /transl_table=11 /product="hypothetical protein" /protein_id="AAO18197.1" /db xref="GI:27657766"
QY	321	MetValAlaIleIle-----ProAlaArgSer	329	/translation="MGRNATANLQVSTRQIDNI SKALNIDWGGDHVAKANARABQL QSAKLEADKPKRIASLELDSVLKTDCTTPGERRDQIMIAAISDKFLKLTSTGAPQ EVAGQADAAALDRFMYATKAAVHSGSPDNPNDH"
Db	7673	CTGGCGGTCTTCCTGGCGGACCCCTGGCACCAGCACATGCGCCGCGCAAGGCGCGCTCC	7732	complement (4112. .4519) /gene="orf6" complement (4112. .4519) /gene="orf6" /codon_start=1 /evidence=not experimental /transl_table=11 /product="hypothetical protein" /protein_id="AAO18198.1" /db xref="GI:27657767"
QY	330	TrrpMetIleArgThrCysAlaProTrrpValPheGlyGluIleLeuGlyLeuValGlu	349	
Db	7733	CTCCAGCTGCTCAAGCGCGCGCGAGAGCGCGCCGACCAAGTCCGTGCGCTCGTGAG	7792	
QY	350	Ala---GlyAlaThrTrpGluAsnLeuArgAspHisLeuHisTyrArgLeuProAlaAla	368	
Db	7793	TCCGCGGCGCGTGGTGGTGG-----GGACGC	7816	
QY	369	AspValArgProPro	373	
Db	7817	GACGTCCTTCGCGCT	7831	
RESULT 12				
AY172684/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
1. .4603				
/organism="Corynebacterium glutamicum"				
/mol_type="genomic DNA"				
/strain="22220"				
/db_xref="taxon:1718"				
/plasmid="pAG3"				
complement (200. .397)				
/note="orf1"				
/codon_start=1				
/evidence=not experimental				
/transl_table=11				
/product="hypothetical protein"				
/protein_id="AAO18193.1"				
CDS				
ORIGIN				

|||||
::: |||||
AAAAACAGTACACAAAGCGTGATCATGCCCTATGG 1030

gene

```
IRGW"
1963. .2754
/locus_tag="MAP0578"
```

gene

[illegible]

Alignment Scores:
Pred. No.:
Score:

Percent Similarity:	40.94%	Conservative:	58
Best Local Similarity:	27.96%	Mismatches:	136
Query Match:	20.40%	Indels:	128
DB:	1	Gaps:	18
US-10-007-527A-2 (1-379) x AE017229 (1-300717)			
Qy	22	AspLysArgGlyIleAtrGHisGluLeuArgProLysLeuGlnGlnIleThrSerGlu	41
Db	275559	GAGCGGCGCGATTCGTTGGGGCGCAAGGCGGATGCTGTGGCAGGGCGTCTTCGTTGAAG	275618
Qy	42	ThrPheAsnAlaCysGlyArg	52
Db	275619	CGCGTGGCGTCTGCTGGCGGGTGTTCATCAACCATGCGGTGGCGATCTCGATGACGGT	275678
Qy	53	ValAsnGlyValThrIle	68
Db	275679	---CAGGGCGGTGGTATTAAGCGTCGGGAGTTCATGACGCGATGGTGGCGAGTCTGCAT	275735
Qy	69	GlyLeuArgSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAla	88
Db	275736	GGTCTGATGACTTCGGGTTCGGTGGCATGGCCGCTTGTCTGGCAGTCATAGCGAAT	275795
Qy	89	HisArgAlaAspGluIleSerGlnVal	106
Db	275796	ACCAGGGCGCGGAATCGGTGGCAGATGAGAGAGTGTATCGTCTGGT---	275852
Qy	107	ValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrp	126
Db	275853	GTGTATCTGTGATGCTTGCAGATGCGCATGCGGAGATGCTTGTCTGACCTCTGG	275912
Qy	127	ThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArg	146
Db	275913	GATTCGTTTCCAGCGCTCGGTTCGGTTCGGAACCTGTAATTGGACGGGCGAGAA	275972
Qy	146	-----	146
Db	275973	GAACGGATGGTGACGCGAGCGGTGGTGGCCCTTCTCCGGAGATCATGGGTGATGCC	276032
Qy	147	GluMetTrpGlyCysAspGlyTrpValArgAlaValGluIleThrHisGlyLys	164
Db	276033	GAGCGTTCGACATAGCTGCGGTCAACCGGTTGTGGAAGCGACTTACGGCAAGCCGAA	276092
Qy	165	-----AsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal	181
Db	276093	CTTGTGGGATGCTGGTGGACCTTCATATCCATGCCCTCGTGTCTCAGTGACCGCTTG	276152
Qy	182	SerGluAsnIleLeuGluSer	188
Db	276153	TCGAGCGGTCTGATTGAGGCGATTGAGCGACTCTGGCGCGGAGTAAATCATGATTGG	276212
Qy	189	-----PheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSer	203
Db	276213	TTGGCGCGCAACGCTCTTTGCTGCTCGAATACATACAGCTGTGTCGAGGATTTGGCAAG	276272
Qy	204	LeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGly	223
Db	276273	GCTGGCTGTGACATGCGG-----GGTGGTGGCGGTTGATGTCGAGAGATTGACGAC	276326
Qy	224	GluAlaAspGlnValLeuAlaLayrLeuThrLys	239
Db	276327	GAGGCGCTGAATACGTTGGCGCATACCTGTCTCAAGGCTACATATCATGTCGGGCAACG	276386
Qy	240	ValGlyMetGluValGlySerGly	255
Db	276387	ATAGGTCCTTGGAGCTGCGCTTCAACGAAAGAGCTCGGCAAGAGCT-----	276440
Qy	256	AsnArgAlaProTrpGluIle	268
Db	276441	AATCAGATCCGTTGCAAGTCTTGGCAATTCGCTGAGTCTGGTATGCTCGCGGTTTC	276500
Qy	269	-----GlyAsp	270
Db	276501	GGAATTCGACCCCGCGTCAATTTGGCTTCTCCCTGCAGGAAACGAGATTGGCTGTA	276560
Qy	271	-----ProGlnAlaLeuGluLeuTrp	277
Db	276561	ATCGACAGTGATACAGCGGAAGTCGCGAGCATCACGGCCCTGGACAGTGGAGAGTATGG	276620
Qy	278	ArgGluPheGluPheGlySerMetGlyArgAlaIleAlaTrpSerArg	294
Db	276621	CATGAGTGGGAGCAAGCGTCTGTGTGTCGTCTGTCAGATTACTTGTCTCGTCGACGTC	276680
Qy	295	-----GlyLeuArgAlaArgAlaGlyLeuGly	303
Db	276681	AACCTGAATCGGCGCGAGATGCTGTGAATGACTTGTGGATGCTGTGGACGTC	276740
Qy	304	AlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProValMetValAla	323
Db	276741	GCAGAACATCGGACGGAAGAAATTCGCGTCATGAGTGCATGCCGAG---TCGTCGGC	276797
Qy	324	IleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpValPheGlyGlu	343
Db	276798	GTTATTAGTCGACAGGTTTGGTACCAAGTGTTCCTGGCTGCCGTCCA-----GGGTG	276848
Qy	344	IleLeuGlyLeuValGluAla	350
Db	276849	ATTGTGATCTCTGGAGGCG	276869
RESULT 14			
E17316			
LOCUS			
Bifidobacterium breve gene for replication protein of pNBb1.			
DEFINITION			
E17316			
ACCESSION			
E17316.1 GI:5711999			
VERSION			
JP 1998262670-A/1.			
KEYWORDS			
Bifidobacterium breve			
SOURCE			
ORGANISM			
Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;			
Bifidobacteriaceae; Bifidobacterium.			
REFERENCE			
1 (bases 1 to 2297)			
AUTHORS			
Iino, T. and Morishita, T.			
TITLE			
SHUTTLE VECTOR FOR BIFIDOBACTERIA AND REPLICATION PROTEIN GENE OF			
BIFIDOBACTERIA PLASMID			
JOURNAL			
Patent: JP 1998262670-A 1 06-OCT-1998;			
YAKULI HONSHA CO LTD			
COMMENT			
OS Bifidobacterium breve			
FN JP 1998262670-A/1			
PD 06-OCT-1998			
PF 27-MAR-1997 JP 1997091387			
PI IINO TORU, MORISHITA TAKASHI			
PC C12N15/09, (C12N15/09, C12R1:19), (C12N15/09, C12R1:01); CC			
strandedness: Double;			
CC topology: Linear;			
FH Key			
FH Location/Qualifiers			
FT source			
FT 1..2297			
FT /organism='Bifidobacterium breve' FT			
FT /strain='ATCC15698'			
FT /plasmid='pNBb1'			
FT -35 signal			
FT 387..392			
FT FT			
FT -10_signal			
FT 410..415			
FT FT			
FT RBS			
FT 525..530			
FT FT			
FT CDS			
FT 539..2023			
FT /product='replication protein of pNBb1'.			
FEATURES			
source			
1..2297			
/organism='Bifidobacterium breve'			
/mol_type='genomic DNA'			
/db_xref='taxon:1685'			
ORIGIN			
Alignment Scores:			
Pred. No.: 4,24e-17 Length: 2297			
Score: 399.50 Matches: 116			
Percent Similarity: 42.25% Conservative: 64			

4 ValSerAlaGluHisLeuSerGly---LysAspArgPro----- 15
650 GTAGCCGCGAAGCGCGAAGCGCGCGGTAGCGCCGCCCTTGGGAAAGACGAAATAA 709
16 -----ProValSerSerAspLysArg-----GlyLeuArgHisGlu 29
710 ATAACCCCTCCCTCTCCGACACCGATTACGGCGCTGGCGTATGCTGCCTCGCTGAA 769
30 LeuArgProLysLeuGlnGlnIleThrThrSerGluThrPheAsn----- 44
770 AGCCGAAAGATTCTCGTCCGTCATCCGCGTGGCGAAACGCTCGGATTCGAGCCGATTAA 829
45 -----AlaCysGlyArgProIleSerGlyValAsnGlyValThrIleVal 59
830 CTGCGCGCTCGCGACGCTGCGCCCAACCGGTGGAC-----ACCGCGTGGGTGTGATG 883
60 AsnGlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyThrIleCys 77
884 ACCAAGCGCGAAGACCGCGTTTACAGCACCATCTGTGCGGCTCGATCTGGGCATCG 943
80 ProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAla 99
944 CCCACCTGCTCGCAATCATTCGCCACGAACCGCCACGAGGTGCGCTCGCCATCGGG 100
100 HisGlnLeu----- 102
1004 AACCATGCGGAAACTGAGAAAGCGCGCGACCAATGCGAGCGAAGCATGAGGGG 106
103 -----GlyThrGlySerValAlaMetValThr 111
1064 CAGCGTCTGCCCGGCACTCATGTTGTCGACAGCTTCGGAACACTACATCTTCGCGACC 112
112 MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAla 131
1124 CTCACCTTCGCGCATCATCGCAATGCGCGCTGGCCATGACCTTGACGCAATCTCTCAAG 118
132 AlaTrpLysAlaAlaThrAsnGlyArgAspTrpArgThrGluArgGluMetTrpGlyCys 151
1184 GGCTGACGAAATGATTAAACGAGCCCTTGGCAACGGCGCTCGGAAAGTGGAAATC 124
152 AspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGlyTrpHisValHisVal 171
1244 AGGGGTTCGTCGCGCGCATTTGAAATTCACCTACGCGTGTGAACGGCTGGCACCCCTCAATT 130
172 HisAlaLeuMetPheSerGlyAspValSerGlyAsnIleLeuGluSerPheSerAsp 191
1304 CATTCGTCATGTTCTCGATGCGCATCTGGACATGGCGAGGTGAGCGCAATGCGACAA 136
192 AlaMetPheAspArgTrpThrSer-----LysLeuValSerLeuGlyPhe----- 206
1364 TGGCTGCTCGATCGCTGGAACCAATGCTCAAGCGCGTTTGCACAGGCATCAAGAAAAA 142
207 -----AlaAlaProLeuArgAsnSerGlyGlyLeuAspValArg--- 219
1424 GAGGCAACCCCTACAACTGCGCCCG---AAGCAGCAACACGCGATAGATCTGCAATTC 148
220 LysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGly 239
1481 AAGTCGGCAAGGACGCGCAACCGCTGCGCGCGAATACATCACCAAGATTCAAGGCGAC 154
240 ValGly-----MetGluValGlySerGlyAspGlyLysSerGlyArgHis 254
1541 AAGGCGCGGTCACTCTGGCTCAGAAATTCGCGCGCGCATCATCAAGATGTCGTATG 160
255 GlyAsnArgAlaProTrpGluIleAlaValAspAlaValGlyGly-----AspProGln 277
1601 GGGTCGTTAAACCGTTTCAATTTGCTGAATCCGGGTGCTCGGGCTGTCGAGTTTCCAG 160

Search completed: October 23, 2004, 05:46:07
Job time : 5476 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 03:58:12 ; Search time 641 Seconds

(without alignments)
3103.792 Million cell updates/sec

Title: US-10-007-527A-2

Perfect score: 2005

Sequence: 1 MTSVSAHSLGKDRPPVLVS.....HLHYRLPAADVRRPIISVRK 379

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10007527/runat_22102004_114115_26609/app_query.fasta_1.519
-DB=N_Geneseq_23Sep04 -QPM=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pt0 -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10007527@cgn_1.1_470 @runat_22102004_114115_26609 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2005	100.0	1140	6 ABQ76122	Rhodococc
C 2	2005	100.0	6334	6 ABQ76124	Rhodococc
C 3	2005	100.0	9652	6 ABQ76126	Plasmid p
C 4	2005	100.0	11241	6 ABQ76125	Plasmid p
C 5	2005	100.0	11241	12 ADH10182	Adh10182 E. coli-R
6	399.5	19.9	2297	2 AAV58945	AAV58945 B. breve

C	7	337	16.8	9431	2	AAQ20907	DNA encod
	8	326	16.3	4447	3	AAA90934	B. lactof
	9	326	16.3	4447	6	ABL49733	Brevibact
	10	326	16.3	4447	6	AAD22582	B. lactof
	11	318	15.9	4447	3	AAA90951	B. lactof
	12	318	15.9	4447	6	ABL49734	Brevibact
	13	318	15.9	4447	6	AAD22583	B. lactof
	14	318	15.9	4447	10	ADB66212	B. lactof
C	15	296	14.8	8500	6	ABA93871	E. coli/c
C	16	282	14.1	2401	6	AAAS17124	Replicon
C	17	282	14.1	4005	6	AAAS17120	Ketogulon
	18	273	13.6	2534	3	AAZ39508	DNA seque
	19	273	13.6	3741	3	AAZ39509	DNA seque
C	20	214	10.7	5648	3	AAZ46134	Endogenou
C	21	133	6.6	349980	6	ABQ81845	Sequence
C	22	127	6.2	3331	1	AAAN91248	Seqence
C	23	125	6.2	349980	6	ABQ81844	Bifidobac
C	24	125	6.2	349980	6	ABQ81844	Bifidobac
C	25	125	6.2	349980	6	ABQ81842	Bifidobac
C	26	125	6.2	349980	6	ABQ81842	Bifidobac
C	27	122.5	6.1	110000	4	AAI99682_02	Continuation (3 of
	28	122	6.1	1686	11	ABD01551	Pseudomon
C	29	122	6.1	1803	11	ABD01566	Pseudomon
C	30	122	6.1	2610	11	ABD01571	Pseudomon
C	31	122	6.1	2871	4	AAAS4190	Pseudomon
C	32	122	6.1	2871	8	ACA42440	Prokaryot
C	33	122	6.1	3790	10	ADC39175	Novel hum
C	34	122	6.1	5054	10	ADF74204	Human nov
C	35	122	6.1	5236	10	ADC30279	Human nov
C	36	122	6.1	5372	12	ADQ25367	Human nov
C	37	120.5	6.0	110000	4	AAI99683_02	Continuation (3 of
C	38	119	5.9	1428	11	ABD09766	Pseudomon
C	39	119	5.9	1910	2	AAV41758	Nitrosomo
C	40	119	5.9	2427	8	ACA37583	Prokaryot
	41	119	5.9	3528	11	ABD09869	Pseudomon
	42	117	5.9	135638	10	ABX34289	S. atrool
	43	117	5.8	68750	3	AAZ55887	Sorangium
C	44	116	5.8	1542	6	AAD26199	Corynebac
C	45	116	5.8	4466	6	ABN59752	Novel hum

ALIGNMENTS

RESULT 1
ABQ76122
ID ABQ76122 standard; DNA; 1140 BP.
XX
AC ABQ76122;
XX
DT 13-JAN-2003 (first entry)
XX
DE Rhodococcus AN12 replication protein Rep DNA.
XX

XX Plasmid stability protein; replication protein; ethylene forming enzyme;
XX carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
XX polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
XX alcohol dehydrogenase; terpene synthase; cholesterol oxidase; gene;
XX shuttle vector; Rep; ds.
XX Rhodococcus erythropolis.
XX WO200255709-A2.

XX 18-JUL-2002.
XX 12-DEC-2001; 2001WO-US047868.
XX 12-DEC-2000; 2000US-0254868P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Bramucci MG, Cheng Q, Kostichka KN, Tomb J;
XX

DR WPI; 2002-557827/59.
DR P-PSDB; ABB84278.
XX
PT New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.
XX
PS Claim 2; Page 63-64; 96pp; English.

XX This invention describes a novel nucleic acid encoding a replication
CC protein or a plasmid stability protein. The product of the invention is
CC useful for expression of nucleic acid such as genes encoding enzymes
CC involved in the production of isoprenoid molecules, polyhydroxyalkanoic
CC acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile
CC hydratases, ethylene forming enzyme, pyruvate decarboxylase, an alcohol
CC dehydrogenase, terpene synthases, and cholesterol oxidase in an
CC Actinomycetales bacteria. The replication protein or plasmid stability
CC protein are useful in cloning and expression vectors and particularly in
CC shuttle vectors for the expression of homologous and heterologous genes
CC in Rhodococcus sp. This sequence encodes the Rhodococcus AN12 Rep protein
CC described in the disclosure of the invention

XX Sequence 1140 BP; 241 A; 287 C; 347 G; 265 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,41e-157 Length: 1140
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-007-527A-2 (1-379) x ABQ76122 (1-1140)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
DB 1 ATGACCAGCGTAAGTGTGAACACTTTCGGGCAAGACCGGCTCCGCTCGTGTGCG 60
QY 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
DB 61 TCCGATAAGCGCGGCATCCGGCACGAACTGGCGACCCAACTTCAACAAATCACCACGTC 120
QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
DB 121 GAAACATTTAAACGCTGTGGCGCGCGAATTTCTGGCGTGAACGGGTGACCATTTGTCAAC 180
QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyLysGlyTrpIleCysPro 80
DB 181 GGTCCGAAAGGTCTCGATTTCGGAGCCCTTCGTTCTCGGAAAGGGCTGGATCTGCCCC 240
QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
DB 241 TGCTGTGGGGAAGTCCGTGCACATCGTCGACAGCAAAATTTCTCAAGTTGTGCTCAT 300
QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
DB 301 CAATTCGGGACTGGATCTGTTCCGATGTGTGACGATGACCATCGCCATACAGCTGGTCAG 360
QY 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
DB 361 CGGCTCCAGCACTATGACTGGACTTTTCGGCAGCCTCGAAAGCTGGCAACACGGTCGT 420
QY 141 ArgTrpArgThrGluArgGluMetThrGlyCysAspGlyThrValArgAlaValGluIle 160
DB 421 CGTTGGCGTACGGAACGTGAATGTACGCTCGCAGCGATACGTGCGGCTGTTGAATC 480
QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
DB 481 ACTCAGGAAACCGCTGGCACGCTCCACGTTCCAGCGCTACTCATGTTCACTGGTGCAC 540
QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
DB 541 GTGAGTGAACAATCTCGAATCTTCTCGATGCGATGCGATGCGTGGAGTCTCCAAA 600

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
DB 601 CTCGTATCTCTGGATTTGCTGCGCCACTACGTAATTCGGGTGGTCTCGATGTACGAAG 660
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrlleuThrLysIleAlaSerGlyVal 240
DB 661 ATCGCGGCTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCATCTGGCGTT 720
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
DB 721 GGTATGAGGTTGCTAGTGGCGACGGGAAAGTGGTCGACATGGCAACCTGGACCTGG 780
QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
DB 781 GAAATCTCTGTTGATGAGTGGCGGGGATCCACAGCGTTGGAACTGTGGCGAATTT 840
QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAla 300
DB 841 GAGTTTGGTTCGATGGGACGTCGGCAATCGGTGGTCCGTTGGATTGGCTGCCGAGCT 900
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
DB 901 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGATCTGCCCGGTC 960
QY 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
DB 961 ATGTTTGGATCATTCGCGCGGATCGTGGATGATGATTCGACTTGTGGCCTTACGTC 1020
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
DB 1021 TTGCGCGAGATCCTCGGACTCGTCAAGCTGGCGGACTTGGGAAAAATCTTCGTGATCAC 1080
QY 361 LeuHisTrpArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
DB 1081 TTGCATTTATCGATTGCCCGCAGCGGATGTGCGGCCCCCGATAATATCGTTCCGAAG 1137
RESULT 2
ABQ76124/c
ID ABQ76124 standard; DNA; 6334 BP.
XX
AC ABQ76124;
XX
XX 13-JAN-2003 (first entry)
DE Rhodococcus AN12 derived plasmid pAN12 DNA.
XX
XX Plasmid stability protein; replication protein; ethylene forming enzyme;
KW carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;
KW polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase;
KW alcohol dehydrogenase; terpene synthase; cholesterol oxidase;
KW shuttle vector; circular; ds.
XX
OS Rhodococcus erythropolis.
OS Synthetic.
XX
XX WO200255709-A2.
FN
XX 18-JUL-2002.
PD
XX 12-DEC-2001; 2001WO-US047868.
PF
XX 12-DEC-2000; 2000US-0254868P.
PR
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA
XX Bramucci MG, Cheng Q, Kostichka XN, Tomb J;
PI
XX WPI; 2002-557827/59.
DR
XX New nucleic acid molecule encoding replication protein/plasmid stability
PT protein, useful in cloning and expression vectors, particularly shuttle
PT vectors for expression of heterologous genes in Rhodococcus species.

XX Claim 19; Page 68-71; 96pp; English.

XX This invention describes a novel nucleic acid encoding a replication protein or a plasmid stability protein. The product of the invention is useful for expression of nucleic acid such as genes encoding enzymes involved in the production of isoprenoid molecules, polyhydroxyalkanoic acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol dehydrogenase, terpene synthases, and cholesterol oxidase in an Actinomycetales bacteria. The replication protein or plasmid stability protein are useful in cloning and expression vectors and particularly in shuttle vectors for the expression of homologous and heterologous genes in *Rhodococcus* sp. This sequence represents the *Rhodococcus* AN12 derived CC plasmid pAN12 DNA described in the disclosure of the invention XX

SQ Sequence 6334 BP; 1386 A; 1807 C; 1681 G; 1460 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.49e-156	Length:	6334
Score:	2005.00	Matches:	379
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-007-527A-2 (1-379) x ABQ76124 (1-6334)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
 DB 3051 ATGACCGGTAAGTGTGAACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTG 2992

QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrSer 40
 DB 2991 TCCGATAAGCGCGCATCCGCGCAAGACTCGACCCCAACTTCAACAATCACCACGTC 2932

QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
 DB 2931 GAAACATTTAACGCTGTGGCGCGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTCA 2872

QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrPileCysPro 80
 DB 2871 GTCCCGAAGGTTCTGGATTGGAGGCTTCTGTTCTCGGAAAGGGTGGATCGCC 2812

QY 81 CysCysAlaGlyLysValGlyValHisArgAlaAspGluIleSerGlnValValHis 100
 DB 2811 TGCTGTCGGGAAAGTCGGTGCACATCGTGCACGAAATTTCTCAAGTTGTTCAT 2752

QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
 DB 2751 CAATCGGGACTGGATCTGTTGGATGGTGCATGACCATGACCATGCGCCATACAGTGGTCAG 2692

QY 121 ArgLeuHisAspLeuThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140
 DB 2691 CGGCTCCACGACCTATGGATCGAATGTACGGTGGACGATCGGAAAGCTGCGCAACGGTCTG 2632

QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 DB 2631 CGTTGGCGTACGAACTGTAATGTACGGTGGACGATACGATCGCGCTGTGTGAATC 2572

QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 DB 2571 ACTCAGGAAACGCGCTGGCAGCTCCAGCTTCCAGCGCTACTCATGTTCACTGGTGAC 2512

QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
 DB 2511 GTGAGTGAGAACATCTCGAATCTTCTCGGATCGCATGTTCCATCGGTGACTTCCAAA 2452

QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLeuAspValArgLys 220
 DB 2451 CTCGTATCTCGGATTTCTGGCGCCACATACGTAATTCGGGTGGTCTCGATGTACGAAA 2392

QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240

DB 2391 ATCGCGGGGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTCGATCTGGCTT 2332

QY 241 GlyMetGluValGlySerGlyLysSerGlyArgHisGlyValAsnArgAlaProTrp 260

DB 2331 GGTATGGAGTTGGTAGTGGCGACGGAAGAAAGTGGTGGACATGGCAACCGTGACCTGG 2272

QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280

DB 2271 GAAATCGCTGTGTGATGCGTGGCGGGATCCACAAGCGTTTGAACCTGTGGCGAGATTT 2212

QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300

DB 2211 GAGTTTGGTTGATGGGACGTCGGCAATCGCGTGGTCCGCTGGATTGGTCCCGAGCT 2152

QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320

DB 2151 GGTCTTGGCGCAGAACTAACAGATGCTCAGATCGTTGAGCAGAAAGATCTGCCCGGTC 2092

QY 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal 340

DB 2091 ATGGTTGCGATCATTTCCGGCGCATCGTGGATGATTCGACTTGTGGCTTACGTC 2032

QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360

DB 2031 TTCGGCAGATCTCTCGGACTCTCGAAGCTCGCGGACTTGGGAAATCTTCGTGATCAC 1972

QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379

DB 1971 TTGCATTATCATTTGCCCGCAGCGGATGTGCGCCCCCGATAATATCGGTTGCGCAAG 1915

RESULT 3

ABQ76126/c

ID ABQ76126 standard; DNA; 9652 BP.

XX

AC ABQ76126;

XX

DT 13-JAN-2003 (first entry)

XX

DE Plasmid pRHB171 DNA.

XX

KW Plasmid stability protein; replication protein; ethylene forming enzyme; carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase; polyhydroxyalkanoic acid synthase; PHA synthase; nitrile hydratase; alcohol dehydrogenase; terpene synthase; cholesterol oxidase; shuttle vector; circular; ds.

XX

OS Synthetic.

XX

PN WO200255709-A2.

XX

PD 18-JUL-2002.

XX

PF 12-DEC-2001; 2001WO-US047868.

XX

PR 12-DEC-2000; 2000US-0254868P.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI Bramucci MG, Cheng Q, Kostichka KN, Tomb J;

XX

DR WPI; 2002-557827/59.

XX

PT New nucleic acid molecule encoding replication protein/plasmid stability protein, useful in cloning and expression vectors, particularly shuttle vectors for expression of heterologous genes in *Rhodococcus* species.

XX

PS Claim 27; Page 72; 96pp; English.

XX

CC This invention describes a novel nucleic acid encoding a replication protein or a plasmid stability protein. The product of the invention is useful for expression of nucleic acid such as genes encoding enzymes involved in the production of isoprenoid molecules, polyhydroxyalkanoic

QY	261	GlulIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGlulLeuTrpArgGluPhe	280
Db	4272	GAATTCGCTGTTGATCGAGTCAGTCGGCGGGATCCACAAGCGTTTGGAACTGTGGCGAGAATTT	4213
QY	281	GlupheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300
Db	4212	GAGTTTGGTTTCGATCGGACGTCGCGCAATTCGGTGTGTCCTCCGTGGATTTGGTGTCCCGAGCT	4153
QY	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluGluSerAlaProVal	320
Db	4152	GGTCTTGTTGGGCGAAGAACTAACAGATGCTCAGATCGTTTCAGCAGGAGAGAATCTGCCCGGTC	4093
QY	321	MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal	340
Db	4092	ATGTTTGGCATTCATTCGCGCGCATCGTGGATGATGATTTCGGACTTGTGGCGCTTACGTC	4033
QY	341	PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTroGluAsnLeuArgAspHis	360
Db	4032	TTGCGCGAGATCCTCGGATCTCGTTCGAAGCTGGCGGCACATTGGGAAAATCTTCGTGATCAC	3973
QY	361	LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys	379
Db	3972	TTGCATTATCGATTGCCGCGCGGATGTGGGCCCCCGATAATATATCGGTTGCAAG	3916
RESULT 4			
ID	ABQ76125/c		
XX	ABQ76125 standard; DNA; 11241 BP.		
AC	ABQ76125;		
XX	DT		
XX	13-JAN-2003 (first entry)		
XX	Plasmid pRHBR17 DNA.		
DE			
XX	Plasmid stability protein; replication protein; ethylene forming enzyme;		
KW	carotenoid biosynthesis enzyme; isoprenoid; pyruvate decarboxylase;		
KW	polynhydroxalkanoic acid synthase; PHA synthase; nitrile hydratase;		
KW	alcohol dehydrogenase; terpene synthase; cholesterol oxidase;		
KW	shuttle vector; circular; ds.		
XX			
OS	Synthetic.		
XX			
XX	WO200255709-A2.		
XX			
XX	18-JUL-2002.		
XX			
PF	12-DEC-2001; 2001WO-US047868.		
XX			
PR	12-DEC-2000; 2000US-0254868P.		
XX			
PA	(DUPO) DU PONT DE NEMOURS & CO E I.		
XX			
PI	Bramucci MG, Cheng Q, Kostichka KN, Tomb J;		
XX			
DR	WPI; 2002-557827/59.		
XX			
XX	New nucleic acid molecule encoding replication protein/plasmid stability		
PT	protein, useful in cloning and expression vectors, particularly shuttle		
PT	vectors for expression of heterologous genes in Rhodococcus species.		
XX			
PS	Claim 26; Page 71-72; 96pp; English.		
XX			
CC	This invention describes a novel nucleic acid encoding a replication		
CC	protein or a plasmid stability protein. The product of the invention is		
CC	involved for expression of nucleic acid such as genes encoding enzymes		
CC	used in the production of isoprenoid molecules, polyhydroxyalkanoic		
CC	acid (PHA) synthases, carotenoid biosynthesis enzymes, nitrile		
CC	hydratases, ethylene forming enzyme, pyruvate decarboxylase, alcohol		
CC	dehydrogenase, terpene synthases, and cholesterol oxidase in an		
CC	actinomycetales bacteria. The replication protein or plasmid stability		
CC	protein are useful in cloning and expression vectors and particularly in		
CC	shuttle vectors for the expression of homologous and heterologous genes		
CC	in Rhodococcus sp. This sequence represents the plasmid pRHBR17 DNA		

CC described in the disclosure of the invention

XX Sequence 11241 BP; 2517 A; 3109 C; 2945 G; 2670 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.98e-156 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-007-527A-2 (1-379) x ABQ76125 (1-11241)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
DB 6641 ATGACCAACGCGTAAGTGTGAACACCTTTCCGGCAAGACCGCGCTCCGTCGTGTCG 6582
QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
DB 6581 TCCGATTAAGCGCTGTGGCCGCGCATTTCTCGCGTGAACGCTGTGACCATTTGTCAAC 6522
QY 41 GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn 60
DB 6521 GAAACATTTAAGCGCTGTGGCCGCGCATTTCTCGCGTGAACGCTGTGACCATTTGTCAAC 6462
QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
DB 6461 GGTCCGAAGAGTTCTGGATTTCGGAGGCTTCTGTTCTCGGAAAGGCTGATCTGCCCC 6402
QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
DB 6401 TGCTGTGGGAAAGTCGGTGCACATCGTGACACGAAATTTCTCAAGTTTGTGCTCAT 6342
QY 101 GluLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
DB 6341 CAACTCGGGACTGGATCTGTTGCATGCTGACGATGACCATCGCCATACAGCTGGTCAG 6282
QY 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
DB 6281 CGGCTCCACGACCTATGGAATGGAATGTCAGTGCACGATGACGCGCTGTTGAATC 6222
QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTrpValArgAlaValGluIle 160
DB 6221 CGTTGGGTACGAGACGTAATGTCAGTGCACGATGACGCGCTGTTGAATC 6162
QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
DB 6161 ACTCACGAAAAACGCTGGCAGCTCCAGCTTCCAGCGCTACTCATGTTTCAGTGGTGAC 6102
QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
DB 6101 GTGAGTGAGAACATCTCGAATCTCTCGGATGCGATGTCGATCGGCTGCTCCAA 6042
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
DB 6041 CTGATATCTCGGATTGTCGCCACTAGTAATTCGGTGGTCTCGATGACGAAAG 5982
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyVal 240
DB 5981 ATCGCGGTGAAGCTGATCAAGTTCTCGCTCGTATCTGACGAAAAATTTGATCTGGCGTT 5922
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
DB 5921 GGTATGAGGTGGTATGTCGCGACGGAAGAAAGTGTGCATGACACCGTGCACCCCTGG 5862
QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
DB 5861 GAAATCGCTGTTGATGACGTGGCGGATCCACACGCTTGAACCTGTGGCGAATTT 5802
QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
DB 5801 GAGTTTGGTTCGATGGGACGTCGGGCAATCGCGTGGTCCCGTGGATTGGTGGCCGAGCT 5742

QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
DB 5741 GGTCTTGGGGCAGAACTAAACAGATGCTCAGATCTGTGAGCAGGAAGAAATCTGCCCGGTC 5682
QY 321 MetValAlaIleIleProAlaAspSerTrpMetMetIleArgThrCysAlaProTrpVal 340
DB 5681 ATGTTTGGGATCATTCGGCGCGCATCTGGATGATGATTCGGACTTGTGGCGCTTACGTC 5622
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
DB 5621 TTCGGCGGATCTTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 5562
QY 361 LeuHisTrpArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
DB 5561 TTGCATTATGATTCGCCGACGATGTGCGGCCCGCATATATCGTTTCGCAAG 5505
RESULT 5
ADH10182/C
ID ADH10182 standard; DNA; 11241 BP.
XX
AC ADH10182;
XX
DT 11-MAR-2004 (first entry)
XX
DE E. coli-Rhodococcus shuttle plasmid pRHBR17.
XX
KW Aryl carotenoid; beta-ionone; carotene desaturase; CrtU; isorenieratene;
KW chlorobactene; pharmaceutical; food supplement; animal feed additive;
KW food colourant; cosmetic; crtO; ds.
XX
OS Synthetic.
FN WO2003093200-A2.
PD 13-NOV-2003.
PF 06-MAY-2003; 2003WO-US014111.
XX
PR 06-MAY-2002; 2002US-0378312P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
PI Cheng Q, Tao L, Rouviere PE;
XX
DR WPI; 2004-081862/08.
XX
CC Production of aryl carotenoids useful in e.g. pharmaceuticals and food
CC colorants involves bioconversion of cyclic carotenoids comprising beta-
CC ionone ring in the presence of carotene desaturase.
XX
PS Example 3; SEQ ID NO 13; 66pp; English.
XX
CC The invention relates to production of aryl carotenoid compounds. The
CC method involves transforming a host cell comprising a cyclic carotenoid
CC having at least one beta-ionone ring with a gene encoding a carotene
CC desaturase (CrtU) followed by growing the transformed host cell. Also
CC provided is a method for regulating aryl carotenoid biosynthesis in a
CC host cell. The method is useful in the production of aryl carotenoids
CC such as isorenieratene and chlorobactene, which are useful in
CC pharmaceuticals, food supplements, animal feed additives, food colorants
CC and cosmetics. The present sequence represents a E. coli-Rhodococcus
CC shuttle plasmid pRHBR17 nucleotide sequence
XX
SQ Sequence 11241 BP; 2517 A; 3109 C; 2945 G; 2670 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.98e-156 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

```
US-10-007-527A-2 (1-379) x ADH10182 (1-11241)
Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
Db 6641 ATGACCAAGCGTAAAGTGTGAACACTTTCGGGAAAGACCGGCTCCGCTCTCGTGTGC 6582
Qy 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSer 40
Db 6581 TCCGATAAGCGCGCATCGGCAGCAACTGCGAGCCCAAACTTCAACAAATCACCACGTCA 6522
Qy 41 GluThrPheAenAlaCysGlyArgProLysLeuSerGlyValAenGlyValThrIleValAsn 60
Db 6521 GAAACATTTAACGGCTGTGGCCGCGGATTTCTGGCGTGAACGGTGTGACCATGTCAAC 6462
Qy 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db 6461 GGTCCGAAAGTTCTGGATTCCGAGGCTTCTGTTCTCGGAAAGGCTGATCTGCCCC 6402
Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 6401 TGCTGTGGGAAAAGTGGTGCATCGTCAGACGAAATTTCTCAAGTTGTGCTCAT 6342
Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 6341 CAATCGGACTGGATCTGTTCCGATGTGACATGACCATGCCCATACAGCTGGTCAG 6282
Qy 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 6281 CGGCTCCACGACCTATGACTGGACTTTTCGGCAGCTGGAAAGCTGGACCAACGGTCGT 6222
Qy 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 6221 CGTTGGCGTACGGAACGTGAATGTACGGCTGCGACGGATACGTGCGCGCTGTTGAATC 6162
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 6161 ACTCAGGAAAACCGCTGCGACGTCCAGTTTACGGCGTACTCATGTTTCAGTGGTGAC 6102
Qy 181 ValSerGluAenIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db 6101 GTGAGTGAGAACATCTCGAATCTTTCGGATCGCATGTTTCGATCGGTGACATCCAA 6042
Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 6041 CTCGTATCTCTGGGATTTGCTGGCCACTAGCTAATTCGGGTGGTCTCGATGTACGAAAG 5982
Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 5981 ATCGCGGTGAAGCTGATCAAGTTCTCGTGGTATCTGACGAAATTTGCATCTGGCGTT 5922
Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 5921 GGTATGGAGTTGTGTAGTGGCGCGGAAAAGTGTGACATGGCAACCGTGCAACCTGTG 5862
Qy 261 GluIleAlaValAspAlaValGlyCysAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 5861 GAAATCGCTGTGTATGATGAGTGGGCGGGATCCAAAGCGTTGGAACTGTGGCGAAATTT 5802
Qy 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 5801 GAGTTTGTTCGATGGGACGTGGCAATCGCGTGTGTCCTCGTGGATTCGTGGCCGAGCT 5742
Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 5741 GGTCTTGGGCGAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAAATCTGCCCGGTC 5682
Qy 321 MetValAlaIleIleProAlaArgSerTrpMetIleArgThrCysAlaProTrpVal 340
Db 5681 ATGGTTGCGATCATTCGCGCGCATCGTGGATGATTCGAGCTTGTGGCTTACGCTC 5622
Qy 341 PheGlyGluIleLeuLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 5622
```

5621 TTCGGCGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 5562

361 LeuHisTyrArgLeuProAlaAlaAspValArgProPheIleSerValArgLys 379

5561 TTGCATTATCGATTGCCCGCAGCGGATGTGCGGCCCGCATATATCGTTTCGCAAG 5505

RESULT 6

AAV58945

ID AAV58945 standard; DNA; 2297 BP.

XX AAV58945;

XX 05-JAN-1999 (first entry)

XX B. breve essential region gene.

XX Essential region gene; shuttle vector; ds.

XX Bifidobacterium breve.

XX Key Location/Qualifiers

FT -35_signal 387..392 /*tag= a

FT -10_signal 410..415 /*tag= b

FT RBS 525..530 /*tag= c

FT CDS 539..2023 /*tag= d

XX JP10262670-A.

XX 06-OCT-1998.

XX 27-MAR-1997; 97JP-00091387.

XX 27-MAR-1997; 97JP-00091387.

XX (HONS) YAKULT HONSHA KK.

XX WPI; 1998-587288/50.

XX P-PSDB; AAW73071.

XX Shuttle vector for a Bifidobacterium species - contains specific elements from existing plasmids e.g. pNBb1 of ATCC 15698.

XX Disclosure; Page 14-15; 17pp; Japanese.

XX This sequence represents the essential region gene from plasmid pNBb1 used in the vector of the invention. The vector is a shuttle vector for a Bifidobacterium, and contains: (a) a replicated essential region CC originated from a plasmid pNBb1 of Bifidobacterium breve ATCC 15698; (b) a replicated essential region originated from a plasmid of E. coli; and (c) an antibiotic-resistant gene originated from a plasmid of E. coli and an antibiotic-resistant gene functioning by a Bifidobacterium. The CC shuttle vector can be used to transform various Bifidobacteria species

XX Sequence 2297 BP; 524 A; 659 C; 689 G; 425 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.97e-23 Length: 2297

Score: 399.50 Matches: 116

Percent Similarity: 42.25% Conservative: 64

Best Local Similarity: 27.23% Mismatches: 169

Query Match: 19.93% Indels: 77

DB: 2 Gaps: 15

US-10-007-527A-2 (1-379) x AAV58945 (1-2297)

Qy 4 ValSerAlaGluHisLeuSerGly---LysAspArgPro----- 15

Db 650 GTAGCCGCGAAGCGGCAAGCGGCGGTAGCGGCCCGCTTGGGAAAAGACGAAATAA 709

QY	228	ValLeuAlaAlaTyrLeuThrTrpLysIleAlaSerGly-----ValGlyMetGluValGly	245
Db	2913	GACCTCGCGAGTACATCGCAAGACCGAGACGGAGCGGCCCGCCCTCGAACTGCC	2854
QY	246	SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle-----	262
Db	2853	CGCGCGACCTCAAGACGGCGACCGGGGAAACGTGCGCCCGTTCGAACTCCTCGGACGG	2794
QY	263	AlaValAspAlaValGlyGly-----AspProGlnAla-----	273
Db	2793	ATCGGGGACCTGACCGCGCGCATACCGAGGACGACCGCCCGGGGTCGGCTCGCTGGAG	2734
QY	274	-----LeuGluLeuTrpArgGluPheGluPheGlySerMetGlyArgArgAlaIleAla	291
Db	2733	TGGAACCTCTCGCGCTGGCAGCAGTACGACGGGGCAACCCGGGGACCGCGGCCATCGAA	2674
QY	292	TrpSerArgGlyLeuArgAlaAaGAlaGlyLeuGlyAlaGluLeuThrAspAla	309
Db	2673	TGGACCGGTACTCTGGCGCAGATGCTCGGGCTCGACGGCGCGACACCGAGGCC	2620
RESULT	8		
AAA90934			
ID	AAA90934	standard; DNA; 4447 BP.	
XX	AC	AAA90934;	
XX	15-SEP-2003	(revised)	
XX	15-JAN-2001	(first entry)	
XX	B.	lactofermentum p48K coding sequence.	
XX	Temperature sensitive	plasmid; TSRCR; protein production;	
XX	temperature sensitive	replication control region; p48K; ds.	
XX	Corynebacterium	glutamicum.	
XX	Key	Location/Qualifiers	
XX	CDS	1318..2601	
XX	FT	/*tag= a	
XX	EPI038966-A1.		
XX	27-SEP-2000.		
XX	16-MAR-2000;	2000EP-00105326.	
XX	16-MAR-1999;	99JP-00069896.	
XX	(AJIN)	AJINOMOTO CO INC.	
XX	Nakamura J,	Kanno S, Kimura E, Matsui K, Nakamatsu T;	
XX	WPI;	2000-573832/54.	
XX	P-PSDB;	RAY97536.	
XX	Plasmids containing	a temperature sensitive replicative control regions	
XX	useful for breeding	microorganisms for the production of amino acids by	
XX	fermentation.		
XX	Claim 7;	Page 14-17; 29pp; English.	
XX	This sequence	encodes the Brevibacterium lactofermentum p48K protein. The	
XX	invention relates	to a plasmid containing a temperature sensitive	
XX	replication control	region (TSRCR) and a marker gene (MG). The TSRCR is	
XX	derived from	plasmid pAM330 harboured by Brevibacterium lactofermentum	
XX	(ATCC 13869)	and allows the plasmid to replicate autonomously at a low	
XX	temperature but	does not allow the plasmid to replicate autonomously at	
XX	an elevated	temperature in coryneform bacteria within a temperature range	
XX	in which the	bacteria can grow. The plasmid can be used for modifying a	
XX	chromosomal	gene in a coryneform bacterium, which may be used for the	
XX	production of	useful substances, such as amino acids, by fermentation, to	
XX	change their	genetic traits. Therefore, the plasmid can be used for	

2265 GTACAGCTCGCGGTCTCGAAGACCGGATCGAATCAACCGCGTTCCTGTTGC 2324

Db	323	alleleProAlaAArgSerTrpMetMetileArgThr	335
QY	323		
Db	2325	TTTGGTGAAGCCCGATGATTGGAACTGATTCAGTCT	2361
QY	2325		
RESULT 10			
AA022582			
ID	AA022582	standard; DNA; 4447 BP.	
XX	AA022582		
XX	29-AUG-2003	(revised)	
DT	12-FEB-2002	(first entry)	
XX	B. lactofermentum pSF6	replication control region DNA.	
DE			
XX	Coryneform bacteria;	arginine repressor; L-arginine; liver function;	
KW		amino acid infusion; ds.	
XX			
OS	Corynebacterium glutamicum.		
XX			
XX	Key	Location/Qualifiers	
XX	CDS	1318..2598	
FT		/*tag= a	
FT		/product= "B. lactofermentum pSF6 replication control	
FT		region Protein"	
XX			
XX	EP1154020-A2.		
XX	14-NOV-2001.		
PD			
XX	24-APR-2001;	2001EP-00109457.	
PF			
XX	28-APR-2000;	2000JP-00129167.	
PR			
XX	(AJIN) AJINOMOTO KK.		
XX	Suga M, Asakura Y, Mori Y, Ito H, Kurahashi O;		
XX	WPI; 2002-012660/02.		
DR	P-PSDB; AAE13530.		
XX			
XX	Coryneform bacterium with inactivated arginine repressor, useful for		
PT	fermentative production of arginine at high levels.		
PT			
XX	Example 1; Page 12-15; 30pp; English.		
PS			
XX	The present invention relates to Coryneform bacterium in which the		
XX	arginine repressor does not function normally and which produces L-		
CC	arginine. Coryneform bacterium is used to produce L-arginine which is		
CC	useful in agents for promoting liver function and in amino acid infusions		
CC	or in 'comprehensive' amino acid pharmaceutical. Coryneform bacterium		
CC	accumulate L-arginine to higher levels than wild-type strains, e.g. 120		
CC	compared with 20, mg/dl. The present sequence is Brevibacterium		
CC	lactofermentum pSF6 replication control region DNA. This DNA is used in		
CC	the construction of shuttle vector for Escherichia coli and Coryneform		
CC	bacteria and temperature sensitive vector. (Updated on 29-AUG-2003 to		
CC	standardise OS field)		
XX			
SQ	Sequence 4447 BP; 824 A; 997 C; 1378 G; 1248 T; 0 U; 0 Other;		
XX			
Alignment Scores:			
Pred. No.:	5,53e-17	Length:	4447
Score:	326.00	Matches:	101
Percent Similarity:	47.75%	Conservative:	58
Best Local Similarity:	30.33%	Mismatches:	133
Query Match:	16.26%	Indels:	42
DB:	6	Gaps:	14

US-10-007-527A-2 (1-379) x ABL49733 (1-4447)			
QY	34	LeuGlnGlnIleThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal	53
Db	1396	ATGTACAAGATCAACCAATATAGAGCTCTCGCGGGTGCATAGTG-GCGCAGGACGA	1454
QY	54	Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyLeuAr	71
Db	1455	AGCTGTGCGGTGCTCGTGGTCTGCTAACCGT-----GCTTCGAGTTTCAGGCTCTGCA	1508
QY	71	gSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAl	91
Db	1509	AAATCTCTCACTCTCGTGGGGTCCACTCTGGCTGGAATTGGAAGTCATGGCGAACC	1568
QY	91	aAspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAlaMetValth	111
Db	1569	CATTGAGCTGGCTATTGCTACTAAGAATCACTTGGCGGGTGGCGGCTCATGATGT	1628
QY	111	rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe	130
Db	1629	TGTGGCAGCTTTCGACACACCGCTCACAGTCATTTTGGCAGGTGGAAGCGGTATTA	1688
QY	130	rAlaAlaTrpIlysalalaThrAsnGlyArgArgTrpArgThrGluArgGluMetTyGl	150
Db	1689	GACTGCGTACTCTTCGATGGTGGAAACATCTCAGTGAAGAAAGAACGTCACCGTAC	1748
QY	150	yCysAspGlyTyrrValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa	169
Db	1749	GTCGAGACACCTATAGTACTATAGGTACAGACTCTTGGCGGACGTTGGCAGCTT	1808
QY	169	lHisValHisAlaLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh	189
Db	1809	GCACCGCAACATGCTGTTCTTGGATCGTCCACTGTCTCAGCATGAACCTCAAGCG	1868
QY	189	eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr	209
Db	1869	TGAGGATTCATGTTTCCCGCTGCTGCTGCTGTTAAAGCCGGTATGGACGCGCC	1928
QY	209	oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGlyGluAlaAspGl	227
Db	1929	ACTGGTGAGCAGCGGGTCAAACTTGATCAGTGTCTACCTGGGGTGGAGCGTGGCAA	1988
QY	227	nValLeuAlaIleValLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl	247
Db	1989	A---ATGGCAACCTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGACTGG	2036
QY	247	yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa	267
Db	2037	CGCTACTAAACCCGCTCTAAGGGGTGCTACACGCGCTTTCAGATG---TTGGATAT	2093
QY	267	lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl	279
Db	2094	GGCCGATCAAGACGACCGCGGAGATATGACGCTGTTTGGTGGCTCGGTGGCTGA	2153
QY	279	uPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr	299
Db	2154	GTATCAGGTGGTCTTAAACAACTCGATCG---TCCTGGTCACTGGG-----GCTAA	2204
QY	299	gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu--	315
Db	2205	GGTGTCTTGGCATTGATTACATAGACGCTGATGCTGCTGAATGAAGAGAACT	2264
QY	316	-----GluSerAlaProValMetValAl	323

34 LeuGlnGlnIleThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53

```

Db 1396 ATGTACAAGATCACCAATAGTAAGGCTCTGGGGGTGCCTAGGTG-GCGCAGGACGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr 71
Db 1455 AGCTTTGGGGTGTCTGGTGTCTAAACGGT-----GCTTCGACAGTTTGGGGTCTGCA 1508
QY 71 gSerCysGlyLysGlyTyrPilleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91
Db 1509 AAACCTCTCACTCTCGCTGGGGTCACTCTGCCTGAATGGAGTCATGGGGGAACGGCG 1568
QY 91 aAspGluLeuSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
Db 1569 CATTGAGCTGGCTATTGTCTACTAAGATCACTTGGCGGGGTGGCGGCTCATGATGTT 1628
QY 111 rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSe 130
Db 1629 TGTGGGCACTGTTCACACACCGCTCACAGTCATTTGCGCAGGTGAAGCGGGTATTAA 1688
QY 130 rAlaAlaTyrLysAlaAlaThrAsnGlyArgArgTyrArgThrGluArgGluMetTyrGl 150
Db 1689 GACTGGTACTCTTCGATGGTGAAACATCTCAGTGGGAAGAAAGAACGTGCACGGTAGCG 1748
QY 150 YCysAspGlyTyrValArgAlaValGluIleThrHisGly---LysAsnGlyTyrHisVa 169
Db 1749 GGTGGACACACCTATAGTACTATAGGTACAGCTCTTGGCGGAACGGTGGCACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
Db 1809 GCACCGCAACATGCTGTGTCTGGATCGTCCATGCTGTGACGATCAACAGGCGTT 1868
QY 189 eSerAspAlaMetPheAspArgTyrThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 1869 TGAGGATTCCATGTTTCCCGCTGCTGTGCTGTGTTAAGCCGGTATGACGGCC 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGluAlaAspGl 227
Db 1929 ACTCGTGAGCAGCGGGTCAAACTTGATCAGGTGTCTACCTGGGGTGGAGCGTGGCA 1988
QY 227 nValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
Db 1989 A---ATGGCACTTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGACGGCTC 2036
QY 247 YAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyrGluIleAlaValAspAlaVa 267
Db 2037 CGCTACTAAAACCGCGTCAAGGGTGTACACGCCGTTTACAGT---TTGGATATGTT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTyrArgGl 279
Db 2094 GGCCGATCAAAAGCGACCGCGGAGGATATGGAGCTGTTTGTGGCTCGGTGGCTGA 2153
QY 279 uPheGluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaAr 299
Db 2154 GTAAGAGTTGGTGTCTAAACCTCGGTTCG---TCCTGGTCACTGGG-----GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu-- 315
Db 2205 GCGTGCTTTGGCATTGATTACATAGACGCTGATGTAGCTGCTGAATGGAGAAAGACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
Db 2265 GTACAAGCTCGCGGTCTGGAAGCAGCGGAACGGTCAATCAACCGGCTTCTGTTCG 2324
QY 323 alIleIleProAlaArgSerTyrMetMetIleArgThr 335
Db 2325 TTTGGTGAAGCCCGATGATTGGAACATGATTTCAGTCT 2361
RESULT 11
ID AAA90951
XX standard; DNA; 4447 BP.
AC AAA90951;
XX

```

```

DT 15-SEP-2003 (revised)
DT 15-JAN-2001 (first entry)
XX
DE B. lactofermentum pSPK6 coding sequence.
XX
KW Temperature sensitive plasmid; TSSCR; protein production;
KW temperature sensitive replication control region; pSPK6; ds.
XX
OS Corynebacterium glutamicum.
XX

```

```

FH Key Location/Qualifiers
FT CDS 1318..2601
FT /*tag= a
XX

```

```

XX EPI038966-Al.
XX
XX 27-SEP-2000.
XX
XX 16-MAR-2000; 2000EP-00105326.
XX
XX 16-MAR-1999; 99JP-00069896.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Nakamura J, Kanno S, Kimura E, Matsui K, Nakamatsu T;
XX
XX WPI; 2000-573832/54.
XX
XX P-PSDB; AAY97537.
XX

```

Plasmids containing a temperature sensitive replication control regions useful for breeding microorganisms for the production of amino acids by fermentation.

Disclosure; Page 19-22; 29pp; English.

This sequence encodes the Brevibacterium lactofermentum pSPK6 protein. The invention relates to a plasmid containing a temperature sensitive replication control region (TSSCR) and a marker gene (MG). The TSSCR is derived from plasmid pAM330 harboured by Brevibacterium lactofermentum (ATCC 13869) and allows the plasmid to replicate autonomously at a low temperature but does not allow the plasmid to replicate autonomously at an elevated temperature in coryneform bacteria within a temperature range in which the bacteria can grow. The plasmid can be used for modifying a chromosomal gene in a coryneform bacterium, which may be used for the production of useful substances, such as amino acids, by fermentation, to change their genetic traits. Therefore, the plasmid can be used for breeding microorganisms for the production of amino acids by fermentation. The plasmid comprises a TSSCR that allows the plasmid to replicate autonomously at an elevated temperature in coryneform bacteria within a temperature range in which the bacteria can grow. (Updated on 15 -SEP-2003 to standardise OS field)

Sequence 4447 BP; 827 A; 993 C; 1375 G; 1252 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,55e-16	Length:	4447
Score:	318.00	Matches:	100
Percent Similarity:	47.45%	Conservative:	58
Best Local Similarity:	30.03%	Mismatches:	134
Query Match:	15.86%	Indels:	42
DB:	3	Gaps:	14

US-10-007-527A-2 (1-379) x AAA90951 (1-4447)

QY 34 LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53

Db 1396 ATGTACAAGATCACCAATAGTAAGGCTCTGGCGGGTGCATAGGTG-GCGCAGGACGA 1454

QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr 71

Db 1455 AGCTGTGGGGTGTCTGGTGTCTTAAACGGT-----GCTTCGACAGTTTGGGGTCTGCA 1508

QY 71 gSerCysGlyLysGlyTyrPilleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91

DB 1509 AACTCTCTCTCGTGGGGGTCATCTCTGGGTGAATGGAAGTCATGGCGGACCGC 1568
QY 91 aSpGluLeuSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
DB 1569 CATTGAGCTGGCTATTCTCTACTAAGATCATCTTGGCGGGTGGCGGCTCATGATTT 1628
QY 111 rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130
DB 1629 TGTGGGCACTGTTGGACACACCGCTCACAGTCATTTTGGCGAGGTTGAAGCGGTATTA 1688
QY 130 rAlaAlaTrpIysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTrpGl 150
DB 1689 GACTGCGTACTCTCGATGGTGAACATCTCAGTGAAGAAAGACGTCACCGTACGG 1748
QY 150 yCysAspGlyTrpValArgAlaValGluLeuHisGly---LysAsnGlyTrpHisVa 169
DB 1749 GTGGAGACACCTATAGTACTATGAGTACAGCTCTTGGCGGACCGTTCGCACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
DB 1809 GCACCGCAACATGCTGTCTTGGATCGTCCACTCTCTGACATGAACCTCAAGGCATT 1868
QY 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
DB 1869 TGAGGATTCATGTTTCCCGCTGGTCTGGTGTAAAGCCGGTATGGACGCC 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgIysIleGlyGlyGluAlaAspGl 227
DB 1929 ACTGGTGAGCAGCGGGTCAAACTTGATCAGTGCTCTACCTGGGTGGAGCGCTGCGAA 1988
QY 227 nValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
DB 1989 A---ATGCGCAACCTACCTCGCTAG---GGCATGTCTCAGGAACCTGACTGGCTC 2036
QY 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa 267
DB 2037 CGCTACTAAACCGCGTCTAAAGGGTCTACACGCGGTTTCAGATG---TTGGATATGT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu---LeuTrpArgGl 279
DB 2094 GCGCATCAACGACGCGCGGAGGATGATGACGCTGTTTGGTGGTGGTGGCTGGA 2153
QY 279 uPheGluPheGlySerMetGlyArgAlaAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
DB 2154 GTATGAGGTGCTTCTAAACACCTGCTGTCG---TCTTGGTCACTGGG---GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle---ValGluGlnGlu-- 315
DB 2205 GCGTGTGGGCACTGATTTACATAGACGCTGATGCTGTGAATGGAAGAGACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
DB 2265 GTACAGCTCGCGGTCTGGAAGCAGCGGATCGAATCAACCGCGTGTGCTGTGC 2324
QY 323 alleleProAlaArgSerTrpMetMetIleArgThr 335
DB 2325 TTTGTGAAGCCCGATGATTGGAACACTGATTCAGTCT 2361
RESULT 12
ABL49734
ID ABL49734 standard; DNA; 4447 BP.
XX
AC ABL49734;
XX
DE 29-AUG-2003 (revised)
DT 29-MAY-2002 (first entry)
XX
DE Brevibacterium lactofermentum p48K nucleotide sequence SEQ ID NO:7.
XX
KW Kanamycin resistant gene; mutation; high temperature resistance;
KW coryneform microbe; transformation; gene; ds.
XX

Corynebacterium glutamicum.
Key Location/Qualifiers
CDS 1318..2601
/*tag= a
/product= "p48K"
JP2002017362-A.
22-JAN-2002.
05-JUL-2000; 2000JP-00204236.
05-JUL-2000; 2000JP-00204236.
(AJIN) AJINOMOTO KK.
WPI; 2002-263243/31.
P-PSDB; ABB06342.
Transformation of high temperature-resistant coryneform microbe and transformant.
Example; Page 20-22; 28pp; Japanese.
The present invention describes a method for transforming a high temperature-resistant coryneform microbe in which the high temperature-resistant coryneform microbe is treated with a chemical changing the structure of cell wall of the microbe and an electric pulse is applied to a solution containing the microbe having a changed structure of cell wall and a DNA. Also described are: (1) transforming a high temperature-resistant coryneform microbe in which the cell wall of a high temperature-resistant coryneform microbe is weakened and electric pulse is applied to a solution containing the microbe having weakened cell wall and a DNA; (2) a transformant of a high temperature-resistant coryneform microbe in which a recombinant DNA is introduced; and (3) a transformant of a high temperature-resistant coryneform microbe carrying both a plasmid vector derived from Corynebacterium glutamicum or Brevibacterium plasmid vector derived from Corynebacterium glutamicum or Brevibacterium lactofermentum. The method is used for transforming a high temperature-resistant coryneform microbe. The present sequence encodes a Brevibacterium lactofermentum p48K protein sequence, which is used in an example from the present invention. (Updated on 29-AUG-2003 to standardise OS field)
XX Sequence 4447 BP; 827 A; 993 C; 1375 G; 1252 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 2,55e-16 Length: 4447
Score: 318.00 Matches: 100
Percent Similarity: 47.45% Conservative: 58
Best Local Similarity: 30.03% Mismatches: 134
Query Match: 15.86% Indels: 42
DB: 6 Gaps: 14
US-10-007-527A-2 (1-379) x ABL49734 (1-4447)
QY 34 LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
DB 1396 ATGTACAAGATCACCAATAGTAAGGCTCTGGCGGGTGCATAGTG-CGCAGGACGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyLeuAr 71
DB 1455 AGCTGTGGGTGTCCTGCTGCTCAACGGT-----GCTTCGAGTTTGAGGCTCGCA 1508
QY 71 gSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91
DB 1509 AAACCTCACTCTCGTGGGGTCTCTCTGCTGGAATTTGAGTATCATGGCGAACCGCG 1568
QY 91 aaspGluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
DB 1569 CATTGAGCTGGCTATTGCTACTAAGAACTCATTGGCGGGTGGCGGCTCATGATGTT 1628

Sat Oct 23 17:44:52 2004

```

QY 227 GlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySer 246
Db 6520 AAA---ATGCCAACCTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGACTGGC 6473
QY 247 GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAla 266
Db 6472 TCCGCTACTAAACCCGGTCTAAGGGGTCTACACGCCGTTTCAGATG---TTGGATATG 6416
QY 267 ValGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArg 278
Db 6415 TTGCCCGATCAAAAGCAGCGCCGCGAGGATATGACCGTGTTTGGTGGCTCGGTGGCGT 6356
QY 279 GluPheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAla 298
Db 6355 GAGTATGAGGTGGTCTTAAACCTCGGTTTCG---TCCTGGTCACGTGGG-----GCT 6305
QY 299 ArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu 315
Db 6304 AAGCGTGCTTTGGGCATTGATTACATAGACGCTGATGTCGTGTGAATGGAGAGAA 6245
QY 316 -----GluSerAlaProValMetVal 322
Db 6244 CTGTACAAGCTGCCGGTCTGGAAGCACCAGCGGTGCAATCAACCCGGTGTCTGT 6185
QY 323 AlaIleIleProAlaAlaArgSerTrpMetMetIleArgThr 335
Db 6184 GCTTTGGTGAAGCCCGATGATTGGAAACTGATTCACTCT 6146

```

Search completed: October 23, 2004, 04:15:20
Job time : 680 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
C 1	337	16.8	7560	4	US-09-754-112A-2		Sequence 2, Appli
C 2	337	16.8	12738	4	US-09-754-112A-1		Sequence 1, Appli
C 3	326	16.3	4447	3	US-09-521-668B-17		Sequence 17, Appl
C 4	318	15.9	4447	3	US-09-521-668B-19		Sequence 19, Appl
C 5	214	10.7	5648	3	US-09-371-008-1		Sequence 1, Appli
C 6	122.5	6.1	4411529	3	US-09-103-840A-1		Sequence 1, Appl
C 7	122	6.1	1686	4	US-09-252-991A-155		Sequence 155, App
C 8	122	6.1	1803	4	US-09-252-991A-170		Sequence 170, App
C 9	122	6.1	2610	4	US-09-252-991A-175		Sequence 175, App
C 10	120.5	6.0	4403765	3	US-09-103-840A-2		Sequence 2, Appli
C 11	119	5.9	1428	4	US-09-252-991A-8370		Sequence 8370, Ap
C 12	119	5.9	1910	3	US-09-367-051-2		Sequence 2, Appli

[illegible]

```

; FILE REFERENCE: 041144.006
; CURRENT APPLICATION NUMBER: US/09/754.112A
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 12738
; TYPE: DNA
; ORGANISM: Streptomyces lividans
; US-09-754-112A-1

Alignment Scores:
Pred. No.: 8,06e-24 Length: 12738
Score: 337.00 Matches: 107
Percent Similarity: 42.46% Conservative: 45
Best Local Similarity: 29.89% Mismatches: 126
Query Match: 16.81% Indels: 80
DB: 4 Gaps: 16

US-10-007-527A-2 (1-379) x US-09-754-112A-1 (1-12738)
QY 23 LysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSerGluThr 42
Db 1605 CGAAAGCGCGCGGTACAGCGCGCGCGCGGTGTGGGGATCAGCGGGACCGCGG 1546
QY 43 PheAsnAlaCysGlyArg-----ProIleSerGlyVal----- 53
Db 1545 TCGAAGGGCTGGCGCGCGCGCGCTGTGATGACCTCGCTCGCGGTGATCGTCCCGACAGC 1486
QY 54 ---AsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSer 72
Db 1485 GCGGCGGAACGTCTCGTGTCTG-----GGCTGTGTGCGG 1450
QY 73 CysGlyLysGlyTyrIleCysProCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
Db 1449 TCGCGCGGATCTGCTCTGCGCGGTCTGCGCGCGCGACGATCGGCACACGCGCGG 1390
QY 93 GluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAla---MetValThr 111
Db 1389 GAGATCACCGCGCGGTGTGCGAGTGGATCAAGCGCGCGGGACCGCTACCTGTGTCACC 1330
QY 112 MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAla 131
Db 1329 TTCACGCGCGCATGGGCACACGACCGCGCTCGCGGACCTCATGACCGCTCCAGGGC 1270
QY 132 AlaTrpLysAlaAlaThrAsnGlyArg----- 141
Db 1269 ACCCGAAGACGCGCGACAGACCGCGCGCGCGCGCGCTTACCGAGCTGATCACGGGC 1210
QY 142 -----TrpArgThrGlu----- 145
Db 1209 GGCACGTGGCGCGACGCGCGGCAAGGACGCGCGCGCGCGCGCGCGCGCGCGCATC 1150
QY 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLys--- 164
Db 1149 CGAGACCGGATCGGGTACGTCGCGCATGATCGCGCGACCGAAGTACCGCTGGGGCAGATC 1090
QY 165 AsnGlyTrpHisValHisValHisAlaLeuMetPheSerGlyAspVal----- 181
Db 1089 AACGGTGGCACCGCGCATCCACGCGATCCACGCGCGCGCGCGCGCGCGCGCGCGGAG 1030
QY 182 -----SerAspAsnIleLeuGluSerPhe-----SerAspAlaMetPheAspArgTrp 197
Db 1029 CGTCCGCGAAGCAGATCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 970
QY 198 ThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArg----- 211
Db 969 CAGGGGCACCTGGCGTCTCGT-----TGGACCGCGCGCGCGCGCGCGCGCGCGCGCTTC 913
QY 212 -----AsnSerGlyGlyLeuAspValArgLysIleGlyGlyAlaAsp---Gln 227
Db 912 ACGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 853
QY 228 ValIleAlaAlaTrpLeuThrLysIleAlaSerGly-----ValGlyMetGluValGly 245
Db 852 GACCTCGCGAGTACATCGCCAGACCCAGGACGCGGAGCGCGCGCGCGCGCGCGCGCG 793
QY 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle----- 262
Db 792 CGCGCGCGACCTCAAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 733
QY 263 AlaValAspAlaValGlyGly-----AspProGlnAla----- 273
Db 732 ATCGGGGACCTGACCGCGCGCGATGACCGAGAGACGACCGCGCGCGCGCGCGCGCG 673
QY 274 -----LeuGluLeuTrpArgGluPheGluPheGlySerMetGlyArgArgAlaIleAla 291
Db 672 TGGAACTCTCGCGTGGCAGAGTACGAGCGGGCAACCGCGCGCGCGCGCGCGCGCG 613
QY 292 TrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCGCTACCTCGCGCGAGATGCTCGGGCTCGACGCGCGCGCGCGCGCGCGCG 559

RESULT 2
US-09-754-112A-1/c
; Sequence 1, Application US/09754112A
; Patent No. 6692945
; GENERAL INFORMATION:
; APPLICANT: TRIPATHI, Gyanendra, et al.
; TITLE OF INVENTION: Process for the Production of Polyhydroxyoctanoate by Streptomyces
; TITLE OF INVENTION: lividans

```

```

QY 54 ---AsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSer 72
Db 1485 GCGCGCGAAGCTCGGTGTCTG-----GGCTGTATCGG 1450
QY 73 CysGlyLysGlyTyrIleCysProCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
Db 1449 TCGCGCGGATCTGCTCTGCGCGGTCTGCGCGCGCGCGATCCCGGCAAGCGCGCGG 1390
QY 93 GluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAla---MetValThr 111
Db 1389 GAGATCACCGCGCGGTGTGCGAGTGGATCAAGCGCGGGGACCGCTACTGTGTAC 1330
QY 112 MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeuSerAla 131
Db 1329 TTCACGCGCGCGATGGGCACACGACCGCGTCTGCGGACCTCATGAGCGCGCTCCAGGGC 1270
QY 132 AlaTrpLysAlaAlaThrAsnGlyArg----- 141
Db 1269 ACCCGAAGACGCGCGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1210
QY 142 -----TrpArgThrGlu----- 145
Db 1209 GGCACGTGGCGCGACGCGCGGCAAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCATC 1150
QY 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLys--- 164
Db 1149 CGAGACCGGATCGGTACGTGCGCATGATCCGCGCGACCGAAGTACCGTGGGGCAGATC 1090
QY 165 AsnGlyTrpHisValHisValHisAlaLeuMetPheSerGlyAspVal----- 181
Db 1089 AACGGTGGCACCGCGCATCCACGCGATCCGCGCGCGCGCGCGCGCGCGCGCGCGGAG 1030
QY 182 -----SerGluAsnIleLeuGluSerPhe-----SerAspAlaMetPheAspArgTrp 197
Db 1029 CGTCCGCGAAGCAGATCTGCGCCACCTTCGAGCGCGCGCGCGCGCGCGCGCGCG 970
QY 198 ThrSerLysLeuValSerLeuGlyPheAlaProLeuArg----- 211
Db 969 CAGGGGCACCTGGCGGTCTGTG-----TGGACCGCGCGCGCGCGCGCGCGCGCGCTTC 913
QY 212 -----AsnSerGlyLeuAspValArgLysIleGlyGluAlaAsp---Gln 227
Db 912 ACGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 853
QY 228 ValIleAlaAlaTrpLeuThrLysIleAlaSerGly-----ValGlyMetGluValGly 245
Db 852 GACCTCGCGAGTACATCGCCAGACCCAGGACGCGGAGCGCGCGCGCGCGCGCGCGCG 793
QY 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle----- 262
Db 792 CGCGCGCGACCTCAAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 733
QY 263 AlaValAspAlaValGlyGly-----AspProGlnAla----- 273
Db 732 ATCGGGGACCTGACCGCGCGCGATGACCGAGAGACGACCGCGCGCGCGCGCGCGCG 673
QY 274 -----LeuGluLeuTrpArgGluPheGluPheGlySerMetGlyArgArgAlaIleAla 291
Db 672 TGGAACTCTCGCGTGGCAGAGTACGAGCGGGCAACCGCGCGCGCGCGCGCGCGCG 613
QY 292 TrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCGCTACCTCGCGCGAGATGCTCGGGCTCGACGCGCGCGCGCGCGCGCGCG 559

```

```

QY 228 ValLeuAlaAlaTyrLeuThrLysIleAlaSerGly-----ValGlyMetGluValGly 245
Db 852 GACCTCGCCGAGTACATCGCAAGACCGACGAGGAGCGCGCCCTCGAAGCTGCC 793
QY 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle-----262
Db 792 CGGCGCCGACCTCAAGACGCGGACCGCGGGAACGTCGCCCGTTCGAACCTCCTCGGACGG 733
QY 263 AlaValAspAlaValGlyGly-----AspProGlnAla-----273
Db 732 ATCGGGGACCTGACCGCGGCATGACCGAGGACGACGCGCGCGGCTCGCTCGCTGGAG 673
QY 274 -----LeuGluLeuTrpArgGluPheGlySerMetGlyArgAlaIleAla 291
Db 672 TGGAACTCTCGCGCTGCGACGAGTACGAGCGGCAACCGCGGACCGCGGCGCCATCGAA 613
QY 292 TrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCGCTACCTCGCGGAGATGCTCGGGCTCGACGGCGGCGACACCGAGGCC 559

RESULT 3
US-09-521-668B-17
; Sequence 17, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: MATSUI, KAHIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521,668B
; PRIOR FILING DATE: 2000-03-08
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 4447
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1318)..(2598)
US-09-521-668B-17

Alignment Scores:
Pred. No.: 2,36e-23 Length: 4447
Score: 326.00 Matches: 101
Percent Similarity: 47.75% Conservative: 58
Best Local Similarity: 30.33% Mismatches: 133
Query Match: 16.26% Indels: 42
DB: 3 Gaps: 14

US-10-007-527A-2 (1-379) x US-09-521-668B-17 (1-4447)

QY 34 LeuGlnGlnIleThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
Db 1396 ATGTACAGATACCAATAGTAGGCTCTGCGGGGTGCCATAGGTG-GCGCAGGACGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyGlyLeuAr 71
Db 1455 AGCTGTTGCGGTGCTCGTGGTCTTAACGGT-----GCTTCGAGTTGAGGCTCTGCA 1508
QY 71 gSerCysGlyLysGlyTrpIleCysProCysAlaGlyLysValGlyAlaHisArgAl 91
Db 1509 AATCTCTCACTCTCGTGGGGTCCACCTCTGGCTGGAATTGGAAGTCAATGGCGNACGCG 1568
QY 91 aAspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
Db 1569 CATGAGCTGGTATTGCTACTAAGAAATCATCTTGGCGGCGGTGGCGCTCATGATGTT 1628

```

```

QY 111 rMet-----ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130
Db 1629 TGTGGGCACTGTTTCGACACAAACCGCTCACAGTCATTTTCGCGAGTTTGAACGGGTATAA 1688
QY 130 rAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTyrGl 150
Db 1689 GACTGCGTACTCTTCGATGGTGAAACATCTCATGTGGAAGAAAGAAAGCTGCACGGTACGG 1748
QY 150 yCysAspGlyTyrValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa 169
Db 1749 GGTGGAGCACACCTATAGTGCATATGAGGTACAGACTCTTGGCGCAACGGTTGGCACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
Db 1809 GCACGCGCAACATGCTGTTCTTCTTGGATCGTCACCTGTCTGACGATGAACCTCAAGCGGTT 1868
QY 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 1869 TGAGGATTCATGTTTCCGCTGCTGCTGCTGTTGTTAAAGCGCGGTATGACGCGCC 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGluAlaAspGl 227
Db 1929 ACTGCTGAGCACGCGGTCAAACTTGAATCAGTGTCTACTCGGTGGAGCGCTCGGAA 1988
QY 227 nValLeuAlaLalTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
Db 1989 A---ATGGCACCTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGACTGGCTC 2036
QY 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa 267
Db 2037 CGCTACTAAACCGCGCTCTAAGGGGTCTACACGCGCTTTCAGATG---TTGGATATGTT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl 279
Db 2094 GCGCATCAAGCGACGCGCGGAGGATATGACGCTGTTTGTGGCTCGTGGCGTGA 2153
QY 279 uPheGluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
Db 2154 GTATGAGGTGTTCTTAAACACCTCGCTTCG---TCCTGGTCACGTGG-----GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGluIle-----ValGluGlnGlu-- 315
Db 2205 GCGTCTTTGGGCATTTGATTACATAGACGCTGATGCTGTAAGTGAAGAAGAACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
Db 2265 GTACAAGCTCGCGCTCTGGAGCACCGGACCGGTGAAATCAACCGCGTTGCTGTTC 2324
QY 323 alleProAlaArgSerTrpMetMetIleArgThr 335
Db 2325 TTTGGTGAAGCCCGATGATTGGAACCTGATTCACTCT 2361

RESULT 4
US-09-521-668B-19
; Sequence 19, Application US/09521668B
; Patent No. 6303383
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, JUN
; APPLICANT: KANNO, SOHEI
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: MATSUI, KAHIKO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: TEMPERATURE SENSITIVE PLASMID FOR CORYNEFORM BACTERIA
; FILE REFERENCE: 0010-1093-0
; CURRENT APPLICATION NUMBER: US/09/521,668B
; PRIOR FILING DATE: 2000-03-08
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 4447
; TYPE: DNA

```

ORGANISM: Brevibacterium lactofermentum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1318)..(2598)
 US-09-521-668B-19

Alignment Scores:
 Pred. No.: 1.59e-22 Length: 4447
 Score: 318.00 Matches: 100
 Percent Similarity: 47.45% Conservative: 58
 Best Local Similarity: 30.03% Mismatches: 134
 Query Match: 15.86% Indels: 42
 DB: 3 Gaps: 14

US-10-007-527A-2 (1-379) x US-09-521-668B-19 (1-4447)

Qy 34 LeuGlnGlnIleThrThrSerCluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
 Db 1396 ATGTACAGATCACCAATAGTAAGGCTCTGGCGGGTCCATAGTG-GCGCAGGACGA 1454
 Qy 54 Asn-----GlyValThrIleVal-AsnGlyProIleSerGlyPheGlyLeuAr 71
 Db 1455 AGCTTTGGGTTCTCTGCTCTAAGCT-----GCTTCGAGTTTGAGGGTCTGCA 1508
 Qy 71 gSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAl 91
 Db 1509 AACTCTCACTCTCGTGGGGTCTCTCTGCTGAATTGGAAGTCATGGCGGAACGGCG 1568
 Qy 91 aAspGluIleSerGlnValAlaAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
 Db 1629 TGTGGGCACTGTTCGACACACCGCTCACAGTCATTTCGCGAGTTGAAGCGGTATTA 1688
 Qy 130 rAlaIleThrLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgLumetYrGl 150
 Db 1689 GACTGGTACTCTTCGATGGTGAACATCTCAGTGGAGAAAGAACGTCGACGGTACGG 1748
 Qy 150 yCysAspGlyTyTrValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa 169
 Db 1749 GTGGGACACCTATAGTACATAGTACGATGATGACAGACTCTTGGCGGACGGTTGGCATT 1808
 Qy 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
 Db 1809 GCACCGCAACATCTGTGTCTTGGATCGTCCACTGTCTGACGATGAATCAAGGCATT 1868
 Qy 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
 Db 1869 TGAGGATTCATGTTTTCGCGTGGTGTCTGTTGGTTAAGCGCGGTATGACGCGCC 1928
 Qy 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGlyGluAlaAspGl 227
 Db 1929 ACTGCGTGACACGGGTCAAACTTGATCAGGTGTCTACCTGGGTGGAGACGTCGAA 1988
 Qy 227 nValLeuAlaAlaTyreThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
 Db 1989 A---ATGGCAACCTACCTCGCTAAG-----GGCATGTCTCAGGAACCTGAGTGGCTC 2036
 Qy 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa 267
 Db 2037 CGCTACTAAACCGCGTCTAAAGGGTCTGACACCGCGTTTCAGATG---TTGNTATGTT 2093
 Qy 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl 279
 Db 2094 GGCCGATCAAGACGACCGCGCGAGGATATGACGCTGTTTGTGGTCTCGTGGCGTGA 2153
 Qy 279 upheGluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
 Db 2154 GTATGAGGTTGGTCTAAACACCTCGCTTCG---TCTGTGTACGCTGGG-----GCTAA 2204
 Qy 299 sAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu-- 315

Db 2205 GCGTGTCTTGGCGATTGATTACATAGACGCTGATGTACGTCGTAATGGAAGAACT 2264
 Qy 316 -----GluSerAlaProValMetValAl 323
 Db 2265 GTACAAGTCGCGGTCTGGAAGCACCGGACGGGTGATCAACCCCGCTTGTGTGTC 2324
 Qy 323 alIleProAlaAArgSerTrpMetMetIleArgThr 335
 Db 2325 TTTGTGAAGCCCGATGATTGAAACTGATTCACTCT 2361
 RESULT 5
 US-09-371-008-1/c
 ; Sequence 1, Application US/09371008
 ; Patent No. 6127174
 ; GENERAL INFORMATION:
 ; APPLICANT: TONOUCHI, Naoto
 ; APPLICANT: SUGIYAMA, Masakazu
 ; APPLICANT: YOKOZAKI, Kenzo
 ; TITLE OF INVENTION: Plasmid Derived from Gluconobacter Bacteria
 ; TITLE OF INVENTION: and a Vector
 ; FILE REFERENCE: OP873
 ; CURRENT APPLICATION NUMBER: US/09/371,008
 ; CURRENT FILING DATE: 1999-08-10
 ; EARLIER APPLICATION NUMBER: JP 10-227227437
 ; EARLIER FILING DATE: 1998-08-11
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 5648
 ; TYPE: DNA
 ; ORGANISM: Gluconobacter oxydans
 US-09-371-008-1

Alignment Scores:
 Pred. No.: 1.25e-11 Length: 5648
 Score: 214.00 Matches: 91
 Percent Similarity: 40.00% Conservative: 45
 Best Local Similarity: 26.76% Mismatches: 158
 Query Match: 10.67% Indels: 46
 DB: 3 Gaps: 16

US-10-007-527A-2 (1-379) x US-09-371-008-1 (1-5648)

Qy 15 ProProValLeuValSerSerAspLysArgGlyIleArgHisGluLeuArgProLysLeu 34
 Db 3561 CCCCCAGAGCAAAACAGACGCAAAATGCGGTTATCCAGGCTCTTAGAAGGCTTGT 3502
 Qy 35 GlnGlnIleThrThrSerGluThrPheAsnAlaCysGly----- 47
 Db 3501 CACGATCATGACGTGAGATGTTTCGATCCGCGCTTGTGGTTGCTCGCAAGATACTGCC 3442
 Qy 48 -----ArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLys 63
 Db 3441 CACATTTCTCCGCGCTGTGCTGCTGGGATGGATCAACGCGTTG----- 3394
 Qy 64 GlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysProCysAla 83
 Db 3393 -----GCCGTGATGGGTCAAGCATTCGCGCTCTGCT---CGTTGCCCTCACTGTGCG 3343
 Qy 84 GlyLysValGlyAlaHisArgAlaAspGluIleSerGlnVal----- 98
 Db 3342 -----CCTACCGGCATCTGATGTCTGCCCGGTGGTGGCGTCTCTCAAA 3295
 Qy 99 ---AlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThrArgHisThr 117
 Db 3294 GCGGTTATCCCATGATGATACGGCGCT---GCATTCGCCACATGCACTATGCGCATGAC 3238
 Qy 118 AlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThr 137
 Db 3237 CGAAACACTCGCTGTCAAAACATCGCGACGGTTCAGACTAAGGCTATTTCGCGATGCGAG 3178
 Qy 138 AsnGlyArgArgTrp---ArgThrGluArgGluMetTyTrpGlyCysAspGlyTyrValArg 156

```

Db 3177 CGTGGTGGCTTGTGGAACCGGCTGCTCAGGCAC-----GGTGTGCTGGCTTTATCCGG 3124
QY 157 AlavalgluilethrHisGly---LysAsnGlyTrpHisValHisValHisAlaLeuLeu 175
Db 3123 GCTTTGAGTGACTTGGGCGGTAGACGGGCTGGCATCTGCATGCTCCATGCGATTGTT 3064
QY 176 MetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAsp 195
Db 3063 AUTCACCGGACGGCGCAAGGCGCGTT-----AGTGGCGGCGGCTGCCTCAGAGT 3010
QY 196 ArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGly 215
Db 3009 CGCTGGATCAGACTGTGTGCCAAGGCTGTCATATCGCCGTT-----GGA 2965
QY 216 LeuAspValArgLysIleGlyGluAlaAspGlnValLeuAlaAlaTyr---LeuThr 234
Db 2964 TCAGGGCAGAGGTTGTCGGGTATCTGAAGACAAAGGCTTCTGATTATGGGTGCGCA 2905
QY 235 LysIleAlaSer---GlyValGlyMetGluValGlySerGlyAspGlyLysSerGlyArg 253
Db 2904 GATCTTCGAGTTGGGAATGCTTCTGAGATGGCGCGGATGGAAGAAACCGGGAAG 2845
QY 254 HisGlyAsnArg-----AlaProTrpGluIleAlaValAspAlaValGlyGlyAspPro 271
Db 2844 CGCCCTAATCGCCTGAATGTTCTCT---GAGTTGCTTGGCTTGGCGGCTGAAGGTGACCGA 2788
QY 272 GlnAlaLeuLeuLeuTrpArgGluPheGluPheGlySerMetGlyArgAlaIleAla 291
Db 2787 CTTGCTGCCAGAGTACCGGACGACGATTTCTGCCCTGTGCAGTCAACGTTCTGCTGTG 2728
QY 292 TrpSerArgGlyLeuArgAlaAlaArgAlaGlyLeuGlyAlaGlu---LeuThrAspAlaGln 310
Db 2727 GTTGTGCCGATTTGAAGGGATCTTAGATCTGGACTTCGAAGACATTCGACAGAGCAA 2668
QY 311 IleValGluGlnGluSerAlaProValMetValAlaIleIleProAlaArgSerTrp 330
Db 2667 GAGCGGAGCTTGAAGCGCAGCAGAAAGCCTGTGTGGCTTCITCCGGTGTGTTGG 2608

```

RESULT 6

```

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

```

Alignment Scores:
Pred. No.: 614 Length: 4411529
Score: 122.50 Matches: 96
Percent Similarity: 38.19% Conservative: 35
Best Local Similarity: 27.99% Mismatches: 138
Query Match: 6.11% Indels: 76
DB: 3 Gaps: 18

```

US-10-007-527A-2 (1-379) x US-09-103-840A-1 (1-4411529)

QY 46 CysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySer 65

```

Db 224503 TGCACCGACCGCGCGCGCTGTGCACGAAGTTGCGCTGTTGTTCGATCCGACGGGCTC 224444
QY 66 GlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysProCysCysAlaGlyLys 85
Db 224443 GGAGACGGTCACTTGTAGCCTTGGCCCGAGTTGACCGGGACCCCGTA-----CGT 224393
QY 86 ValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHisGlnLeuGlyThr--- 104
Db 224392 CGAGGAGTCCAT-----GACCACCGTGGCGCAACTTCTCAAGCAGCTAGTAGTGCAATT 224339
QY 105 ---GlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHis 123
Db 224338 CGGGTGTGGTGGCGCGGACACCATGCCCCATGACATGGGGAAGGT---CTTTTGCAC 224282
QY 124 AspLeuTrpThrGly---LeuSerAla---AlaTrpLysAlaAlaThrAsnGlyArgArgT 142
Db 224281 GACGCGGTGCGGGTATGTCATCTGATCGCTGCGTCTG-----CGGCGGT 224234
QY 142 TrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArg-AlaValGluIleThr 161
Db 224233 GGCACACCGAATCACCAGTCCGAGCTCGACTTGTGCGCGCGCATCGATGTTTAC 224174
QY 162 HisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal 181
Db 224173 CGCGTGTGGCGGCGCC-AGAACTCAACCGCGCGCAGCTACCTGAGTGGGCTCATCC 224115
QY 182 SerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrp----- 197
Db 224114 AGTAG-AACTTG-----CCGCGCACCGGCGGATGGAAGAAATGTGATGGCGCTTCG 224062
QY 198 -----ThrSerLysLeuValSerLeuGlyPheAlaAlaPro 209
Db 224061 GCCATCGCCCGTGGCGGATCGGCACCGAGTTGATGATCGGTCGGTGGCCACACCG 224002
QY 210 ---LeuArgAsnSerGlyLeuAspVal-----ArgLysIleGlyGluAlaAsp 226
Db 224001 ACCATCGCGCGTTCGACCGGATTGAACGTGGCGCGCGGAAACCGCGCGGCCAACAAAC 223942
QY 227 GlnVal-----LeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGly 241
Db 223941 GATGTCGATCCTCCCGGCAACCGGGCGCGCTCTACCGGCACCTGCGCGGTCAGGC 223882
QY 242 MetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGlu 261
Db 223881 GGGGCGCGCGAGCGGTCAGGTGGCGGCGCGGCCAACCGGTCAGTCCCGCGCGC--- 223825
QY 262 IleAlaValAspAlaValGlyGlyAsp-ProGlnAlaLeuGluLeuTrpArgGluPheG1 281
Db 223824 ---AGCGCCAAACGGGTGGGTGGCGCGCAACA----- 223794
QY 281 uPheGlySerMet-----GlyArgArgAlaIleAlaTrpSerArg----- 294
Db 223793 ---GTCACCTGCGCGGCGGTGCGGTCGCGTGTGTCGCCCGCGCGCGGACCTC 223738
QY 295 -----GlyLeuArgAlaAlaGlyLeuGlyAlaGluLeuThrAspAlaGln-- 310
Db 223737 CGGGATCGCGCGGAGCGGTCGCGGTCCGCGCACCGCAACCGCACCCAGCATCAGCA 223678
QY 311 -----IleValGluGlnGluSerAlaProValMetValAlaIleI1 325
Db 223677 CGCAACCACTGCGGCTATGTTCAAGCGCGCAATAGACTCCCCCTCGTCCAGCCCTGCT 223618
QY 325 eProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrValPheGlyGluIle 345
Db 223617 TCAGCGCGCA-----CATAAATGCTTTTCCC-----TCGGGTGACGTCTCT 223579
QY 345 u 345
Db 223578 T 223578

```

RESULT 7

US-09-252-991A-155

Sat Oct 23 17:44:52 2004

```

; Sequence 155, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 155
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-155

Alignment Scores:
Pred. No.: 0.00686 Length: 1686
Score: 122.00 Matches: 98
Percent Similarity: 33.49% Conservative: 47
Best Local Similarity: 22.63% Mismatches: 166
Query Match: 6.08% Indels: 124
DB: 4 Gaps: 19

US-10-007-527A-2 (1-379) x US-09-252-991A-155 (1-1686)
QY 6 AlaGluHisLeuSerGlyAspArgProValLeuValSerSerAspLysArgGly 25
DB 474 GCTGACGACAACTGCTGGCTCGGCA-----GAAAGCCGAGGA 515
QY 26 IleArgHis-----GluLeuArg---Pro 32
DB 516 CTTGACGACGAACTGCTGGCTGGCGCGCTGTTGAAGAGCGGCTGCTGGAGAA 575
QY 33 LysLeuGlnGlnIleThrSerGluThrPheAsnAlaCysGlyArgProIleSerGly 52
DB 576 CGGCTTCGGCGCGCTCTCAGCAGACCGAAGCTGCTGGCGCGGCTGCTGCTG 635
QY 53 ValAsnGlyValThrIleValIleGlyProLysGlySerGlyPheGlyLeuArgSer 72
DB 636 GTTGACAGCGAACAATTGTTGAAGCGCCCGG-----GCTGCTCTGCTAC 683
QY 73 CysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
DB 684 TCCGGCGCAACCTTCAGCAGACCGGGGCAACTGTT-----CACCGGCT--- 728
QY 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMet 112
DB 729 -----GTTCTGCTGCGCACTGCTGGCGGAGAGCGACCGAGCGCTGCTGAACG 776
QY 113 ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrGlyLeuSerAlaAla 132
DB 777 CTTGCGGCACTGTTGCTGAAG-----GCG 803
QY 133 TrpLysAlaAla-----ThrAsnGlyArgArgTrpArg 143
DB 804 TGGCGGAGAGCTGGTGGCGGCCACAGCAGCGGCTACTCGCGCGCTCGGTGGCGA 863
QY 144 ThrGluArgGluMetTyrGlyCysAspGlyTyrValArg-AlaValGluIleThrHisG1 163
DB 864 -----GTCGCGCGGAGTGGTGGCGGAGCGGAGCGGCTCTGGGCCAA 905
QY 163 YLysAsnGlyTrpHisValHis---ValHisAlaLeuLeuMet-----PheSe 178
DB 906 CCGCGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 965
QY 178 rGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpTh 198
DB 966 CGGCGAGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1025

```

```

QY 198 rSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspVa 218
DB 1026 GAGCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
QY 218 lArGlyIleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLysIleAlaSe 238
DB 1074 CAGCGCGAGCGCGGATCGCGAGGAGCGGCTGCCAGCTTGTGCGCGGCTGCGCGA 1133
QY 238 rGlyValGlyMetGlnValGlySer-----GlyAs 248
DB 1134 GGGCTCGCGAGAGGTGGAAGCGTGGCGGAGAGCGCTTGGCGATGCGCTGCGCGA 1193
QY 248 pGlyLysSerGly-----ArgHisGlyAsnArgAlaProTrpGluI1 262
DB 1194 GCGGAAGAGATCGCGCGCGCGCTGCTGCGCGGAGCGCTGCGAGGATCGC 1253
QY 262 eAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPh 282
DB 1254 CGCTGCGCGGCGATCGCGAGTGGTGGCGCGCATGTCAGGTGGCGAGCGCTTGGAGTT 1313
QY 282 eGlySerMetGlyArgArg---AlaIleAlaTrpSerArgGlyLeuArgAlaArgAlaG1 301
DB 1314 CCAGCGCGTGGCGAGTGGCGCGCCCATGAACGCGAGCGCTGCTGCGGCGCAGCGGCG 1373
QY 301 YLeuGlyAlaGlu-----305
DB 1374 GGTGACGAGGATGCTGCTGCGCGAGCAGCGGTGAAGGCGATCCGCGAGAT 1433
QY 306 -----LeuThrAspAlaGlnIleValGluGlnGluSerAl 318
DB 1434 TTTGCTCACCCACCGGAGATGCTGAGATCGGACTGCGCGAGCGAGGT-GCAACTGC 1492
QY 318 aProValMetValAlaIleIleProAlaArgSer-----TrpMetMetIleArg 334
DB 1493 GCCTGA-ATCGCGCGGAAAGCGCGGCTGAGTGGGTGGTGAAGACACAGCGCG 1551
QY 334 g-----ThrCysAlaProTyrValPheGlyGluLeuGlu 346
DB 1552 GCGCAACAGGAGCGCTGACGACGCGTGTCTGCGCG-----AGCGCGCGCGACCTC 1605
QY 346 YLeuValGluAlaGlyAlaThrTrpGluAsnLeuArg 358
DB 1606 CCGGACCTCGCGCGGCTGCTGCGCGAGCTGCTGCGG 1642

RESULT 8
US-09-252-991A-170/c
; Sequence 170, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 170
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-170

Alignment Scores:
Pred. No.: 0.00756 Length: 1803
Score: 122.00 Matches: 98
Percent Similarity: 33.49% Conservative: 47
Best Local Similarity: 22.63% Mismatches: 166
Query Match: 6.08% Indels: 124

```

```

DB: 4 Gaps: 19
US-10-007-527A-2 (1-379) x US-09-252-991A-170 (1-1803)
QY 6 AlaGluHisLeuSerGlyLysAspArgProProValLeuValSerSerAspLysArgGly 25
D 1273 GCTGGAGCAACTGCTGGCTCGGCCA-----GAACGCCGAGGA 1232
QY 26 IleArgHis-----GluLeuArg---Pro 32
D 1231 CCTCGAGCAACTGGCTGGCGCGCCCTGTTGAAGAGCGCGTGCCTGAGAA 1172
QY 33 LysLeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGly 52
D 1171 CGGCTTCGCGCGCTGCTCCAGCAGACCGAACCGTTGCGCTGGCGACGGCTGCTG 1112
QY 53 ValAsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSer 72
D 1111 GTTGACACGCAACAAATGGTGAAGCGCCCGG-----GCTGGCCTTCGTAC 1064
QY 73 CysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValcIlyAlaHisArgAlaAsp 92
D 1063 TCCGGCGCAACCTGTCAGCAGCACCGGCAACTGCT-----CACCGGGCT--- 1019
QY 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMet 112
D 1018 -----GTTCTGCTGCGCAGCTGCGCGAAGCGCACCGGCTGCTCGAAG 971
QY 113 ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAlaLa 132
D 970 CCTGTGCACCTGTTGCTGGAAG-----GCG 944
QY 133 TrpLysAlaAla-----ThrAsnGlyValArgTrpArg 143
D 943 TGGCGCGAGCTGGTGGCGCCACCAGCAGCGCGAGCTACTCGCGCGCTGGTGGGA 884
QY 144 ThrGluArgGluMetTyGlyCysAspGlyTyValArg-AlaValGluIleThrHisG 163
D 883 -----GCTGCCCGGAGTGGCCCGAGCGCGAGCGGCTCTGGCCAA 842
QY 163 yLysAsnGlyTrpHisValHis---ValHisAlaLeuLeuMet-----PheSe 178
D 841 CCCGATGGCTGATGCTCCCGCGCGCGCGCGCTGGCGAGCTGGCCAAAGGTTTCG 782
QY 178 rGlyAspValSerGluAniIleLeuLeuSerPheSerAspAlaMetPheAspArgTrpTh 198
D 781 CGCGAGATCGGCTACGCTGGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCT 722
QY 198 rSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspVa 218
D 721 GAGCAAGCTGCTGCCCTCGGC-----GCGCGAGCGCGCGAGCGCTGGAGTT 674
QY 218 lArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyThrLysIleAlaSe 238
D 673 CAGCGCGAGCGCGGATCGCGGAGGACGCGCTCCAGCCTGTGTGCGCGCGTGGCGA 614
QY 238 rGlyValGlyMetGluValGlySer-----GlyAs 248
D 613 GGGCTCGCGAGGAGTTCGAGCGCTGGCGGAGGAGCGCTTGCCTCCGCGCGTGGCG 554
QY 248 pGlyLysSerGly-----ArgHisGlyAsnArgAlaProTrpGluI 262
D 553 GCGGAGAGGATGCCCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 494
QY 262 eAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPh 282
D 493 CGGCTCGCGGCGATCGCAGTGGTCCGCGCGCATGTCCAGTGGCGCGAGCGCTTCGAGT 434
QY 282 eGlySerMetGlyArgArg---AlaIleAlaTrpSerArgGlyLeuArgAlaArgAlaG 301
D 433 CCAGCGCGGTGGGAGTTCGCGCGCCCATGAACGCGAGCGCTGCTGCGGCGCAAGCGGC 374
QY 301 yLeuGlyAlaGlu----- 305

```

```

DB 373 GGTGGACGAGAGATCGCTCGGCTGGTTCGAACGACGACGCGTGAAGCGGATCCGCGAGAT 314
QY 306 -----LeuThrAspAlaGlnIleValGlnGlnGluSerAl 318
D 313 TTTCGTCAACCCACCGGAGATGCTCAGCATCCGGAACCTGGCGAGCAGGT-GCAACTGC 255
QY 318 aProValMetValAlaIleIleProAlaArgSer-----TrpMetMetIleAr 334
D 254 GCCTGA-ATCGCGCGGAAAGCGCGGCTGGAGTCGGTGGTGGAGACACGCGG 196
QY 334 g-----ThrCysAlaProTyTrValPheGlyGlyLeuGlu 346
D 195 GCGCAACAGAGCGCGCTGCACGACGCGCTGCTGCGCG-----AGCGCGCGCGCACCTC 142
QY 346 yLeuValGluAlaGlyAlaThrTrpGluAsnLeuArg 358
D 141 CGCGACCTCGCGCGCGGCTGCTGCCAGCTCTGCGG 105

RESULT 9
US-09-252-991A-175/c
; Sequence 175, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 175
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-175

Alignment Scores:
Pred. No.: 0.013 Length: 2610
Score: 122.00 Matches: 98
Percent Similarity: 33.49% Conservative: 47
Best Local Similarity: 22.63% Mismatches: 166
Query Match: 6.08% Indels: 124
DB: 4 Gaps: 19

US-10-007-527A-2 (1-379) x US-09-252-991A-175 (1-2610)
QY 6 AlaGluHisLeuSerGlyLysAspArgProProValLeuValSerSerAspLysArgGly 25
D 2448 GCTGGACGCAACTGCTGGCTCGGCCA-----GAACGCCGAGGA 2407
QY 26 IleArgHis-----GluLeuArg---Pro 32
D 2406 CCTCGAGCAACTGGCTGGCTGGCGCGCGCTGTTGAAGAGCGCGCTGCTGGAGAA 2347
QY 33 LysLeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGly 52
D 2346 CGGCTTCGCGCGCGCTGCTCCAGCAGACCGAACCGTTGCGCGCTGGCGACGCGGTGCTG 2287
QY 53 ValAsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSer 72
D 2286 GTTGACACGCAACAAATGGTGAAGCGCCCGG-----GCTGGCCTTCGTAC 2239
QY 73 CysGlyLysGlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
D 2238 TCCGCGCAACCGTTCAGCAGCACCGGCGCACTGGT-----CACCGGGCT--- 2194
QY 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMet 112
D 1111 -----

```

```

; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 985 Length: 4403765
Score: 120.50 Matches: 96
Percent Similarity: 38.19% Conservative: 35
Best Local Similarity: 27.99% Mismatches: 138
Query Match: 6.01% Indels: 76
DB: 3 Gaps: 18

US-10-007-527A-2 (1-379) x US-09-103-840A-2 (1-4403765)

QY 46 CysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySer 65
DB 224612 TCCACCCAGCACCGCGCTGTGCACGAAGTTCGCTGTTCATCCGACGCGCTC 224553
QY 66 GlyPheGlyGlyLeuArgSerCysGlyLysGlyTyrIleCysProCysCysAlaGlyLys 85
DB 224552 GGAGACGCTCAACTTGTAGCCTTGGCCCGAGTGTACCGGACCCCGTA-----CGT 224502
QY 86 ValGlyAlaHisArgAlaAspGluIleSerGlnValAlaHisGlnLeuGlyThr--- 104
DB 224501 CGAGGAGTCCAT-----GACCACCGTGGCGAACTTCTCAAGCACCTAGTAGTGCATT 224448
QY 105 ---GlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHis 123
DB 224447 CGGGGTCTGTGGCCCGCACCATGCCCATCGACATCGGGAAGGT---CTTTTGAC 224391
QY 124 AspLeuTrpThrGly---LeuSerAla---AlaTrpLysAlaAlaThrAsnGlyArgArg 142
DB 224390 GACGCGTTCGGGTGATTGTCACTGATCGTGGGTGCT-----CGCGGT 224343
QY 142 TrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArg-AlaValGluIleThr 161
DB 224342 GGCCACCGAGATCACCCTCGGAAGCTCGACTTGTGGTCCGCGCATCGATGTTTAC 224283
QY 162 HisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal 181
DB 224282 CGCGGTGTGGCGGCGC-AGAACTCAACCGCGCGCAGCTACTGAGTGGGCTCATCC 224224
QY 182 SerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrp----- 197
DB 224223 AGTAG-AACTTG-----CCGCGACCGCGCGGATGACGAATGTGATGGCGCTTCG 224171
QY 198 -----ThrSerLysLeuValSerLeuGlyPheAlaAPro 209
DB 224170 GCCATCGCCCGTCCGGATCGGCACCGAAGTGTGATGACGATCGCTTGGCCACACCG 224111
QY 210 ---LeuArgAsnSerGlyGlyLeuAspVal-----ArgLysIleGlyGlyGluAlaAsp 226
DB 224110 ACCATCGCGCGTTCGACCGGATTGAACGTCCGGCGCGGACCGCGCGGCGGCCAACAAAC 224051
QY 227 GlnVal-----LeuAlaAlaTrpLeuThrLysIleAlaSerGlyValGly 241
DB 224050 GGTGTGGATCTCCCGCGCAACGGGGCGCGCTCTTACCGCACTGCGCGGCTCAGGC 223991

```

```

DB 2193 -----GTTTCGCTGGCCAGCCTGGCGGAAGCGACCGAGCGGCTGCTCGAACG 2146
QY 113 ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAlaAla 132
DB 2145 CCGTGGACCTGTGTGGAGG-----GCG 2119
QY 133 TrpLysAlaAla-----ThrAsnGlyArgArgTrpArg 143
DB 2118 TGGCGCGAGCTGTGTGGCGCCACCGACGCGCGAGCTACTCGCGCGCTCGTGGCGGA 2059
QY 144 ThrGluArgGluMetTyrGlyCysAspGlyTyrValArg-AlaValGluIleThrHisGln 163
DB 2058 -----GTCGCCCGGAGCTGGCCCGAGCGCGCGAGCGGCTCGTGGCCAA 2017
QY 163 YLysAsnGlyTrpHisValHis---ValHisAlaLeuLeuMet-----PheSer 178
DB 2016 CCGCATGCGCTGATCCCGCGCGCGAGCGCTGGCGAGCTGGCCAGGGTTTGGC 1957
QY 178 rGlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpTh 198
DB 1956 CGGCAGATCCGGGTACGCTGGCCGACAGCGAGCGGCGCCCGCTCTCGCGAAGAGCT 1897
QY 198 rSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspVa 218
DB 1896 GAGCAAGCTGTCTGCCCTCGGC-----CGCGACCGCGCCAGACGCTGGAGTT 1849
QY 218 lArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSe 238
DB 1848 CAGCGCCGAGCGCGATCGCGAGGACGCTGCCAGCTTGTCTGGCGCGTGGCGGA 1789
QY 238 rGlyValGlyMetGluValGlySer-----GlyAs 248
DB 1788 GGGGTTCGCGAGGAGTCAAGCGCTGGCGAGGAGCGCTTCCCGATGCGCTCGCGGA 1729
QY 248 pGlyLysSerGly-----ArgHisGlyAsnArgAlaProTrpIle 262
DB 1728 GCGGAAGAGGATGCCCGCGCGCGCTGGCGCGCGAGCGCTCGACGCGATCGC 1669
QY 262 eAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPh 282
DB 1668 CGCGTCCCGGATCGCGAGTGTCCGCGCATGTCCAGTGGCGCGAGCTTTCAGTT 1609
QY 282 eGlySerMetGlyArgArg---AlaIleAlaTrpSerArgGlyLeuArgAlaArgAlaGl 301
DB 1608 CCAGCGCGCTGGCGAGTCCGCGCCCATGAAACGAGCGGCTGTCTGCGGCGCAAGCGGC 1549
QY 301 YLeuGlyAlaGlu----- 305
DB 1548 GGTGAGCAGGAGATCGTGGCTGTGTCGAACGACGACGCGTGAAGCGATCCGCGAGAT 1489
QY 306 -----LeuThrAspAlaGlnIleValGlnGlnGluSerAl 318
DB 1488 TTTTCGTACCCACCGGAGATGCTCGACGATCCGGAATCGCGAGCAGGT-GCAACTGC 1430
QY 318 aProValMetValAlaIleProAlaArgSer-----TrpMetIleAr 334
DB 1429 GCCTGA-ATCGCGCGGAAACGCGGAGCGGCTGTGAGTGGGTGTGGAAGACAGCGCG 1371
QY 334 q-----ThrCysAlaProTyrValPheGlyGluIleLeuGl 346
DB 1370 GCGCAACAGAGGCGTGCAGGCGCTGCTCGCG-----AGCGCGCGCGGCTC 1317
QY 346 YLeuValGluAlaGlyAlaThrTrpGluAsnLeuArg 358
DB 1316 CGCGACCTCGCGCGGCTGTGCGCAGCTCTCGCG 1280

```

RESULT 10

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.

; ORGANISM: Nitrosomonas sp.
 ; US-09-367-051-2
 Alignment Scores:
 Pred. No.: 0.0168 Length: 1910
 Score: 119.00 Matches: 57
 Percent Similarity: 42.42% Conservatives: 41
 Best Local Similarity: 24.68% Mismatches: 91
 Query Match: 5.94% Indels: 42
 DB: 3 Gaps: 12
 US-10-007-527A-2 (1-379) x US-09-367-051-2 (1-1910)
 QY 71 ArgSerGlyLysGlyThrPheCysProCysCysAlaGlyLysValGlyAlaHisArg 90
 Db 763 CGCTTCGCAAAAGCAACTTCGTGCGCTATGTGCGCTATTCGGCGGGTTCGAAATG 822
 QY 91 AlaAspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAla----- 108
 Db 823 CTT-----ACATCGTATCTTCAACGGTATCAGATCGTTCATCGCTCAATCCGATCTG 876
 QY 109 -----MetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeu 125
 Db 877 AAGCCTTATCTGCTGACTTAAACGGTTAAGAT-----GGACCTGATTTCAGGAACGT 930
 QY 126 TrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgThrGlu 145
 Db 931 TTTGTGCAATTCGCTAAGCTATGCGCTGATGTTCAAAATCTCGCAGAAATACTAATCAA 990
 QY 146 ArgGluMetTyr-----GlyCysAspGlyTyrValArgAlaValGluIleThr 161
 Db 991 GCCCAGCGTTTCGTGAGTTTCTAAGTCTGAGGTTGTTTCATTCGATTCGATGAGTAACT 1050
 QY 162 HisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal 181
 Db 1051 AACCGGGCAACGGCTGGATCTCAGCGCATATGATCTGTTGTGTAGGAGGAAACCA 1110
 QY 182 SerGluAsnIleLeuSerPheSerAspAlaMetPheAspArgTrp----- 197
 Db 1111 GATCAGTCGCTCTTTTCGGCC-----GAATGGCGAGAGATTACC 1149
 QY 198 ---ThrSerLysLeuValSerLeuGlyPheAlaAlaProLeuArgAsn-----SerGly 214
 Db 1150 GGTGATTCACGTGGTGTACGTC-----AGCCCTTAACTGATCCAGTTGACGGC 1200
 QY 215 GlyLeuAspValArgLysIleGly-----GlyGluAlaAspGlnValLeu 229
 Db 1201 TTTCTTGAGTTTGTAAAGTACGCTTAAAGTCTCTGATCTTGAAGCTGGCCATCTTTTC 1260
 QY 230 AlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGlyAspGly 249
 Db 1261 GAAGCGTAC-----AAGATTTTGTCTGTTCTCGCTCGCTC-----GATAGT 1302
 QY 250 LysSerGlyArgHisGlyAsnArgAlaProTyrGluIleAlaValAspAlaValGlyGly 269
 Db 1303 CACGGATGTAGAGAGTTTCGTTACCGGATGATCTTGTGATGATGATGATGATGATGATG 1362
 QY 270 AspProGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 280
 Db 1363 TTGCCG--TTGCTGAATCTTCTATCGATTT 1392
 RESULT 13
 US-09-252-991A-8473
 ; Sequence 8473, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 8473
 ; LENGTH: 3528
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-8473
 Alignment Scores:
 Pred. No.: 0.0413 Length: 3528
 Score: 119.00 Matches: 91
 Percent Similarity: 34.65% Conservatives: 41
 Best Local Similarity: 23.88% Mismatches: 117
 Query Match: 5.94% Indels: 134
 DB: 4 Gaps: 18
 US-10-007-527A-2 (1-379) x US-09-252-991A-8473 (1-3528)
 QY 24 ArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSerGluThrPhe 43
 Db 38 AGAGCGTAATAGTGTGATGCGACCGGAGCCCAACACCTGCCAGATCTGTCGCCGTG 97
 QY 44 Asn-----AlaCysGlyArgProIleSerGlyValAsnGly 55
 Db 98 CGAAATCTGGCAGGCTACCGGATCTGTAGTGTGACACAGCCGCTCGACGGCTTGGCGGC 157
 QY 56 ValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLys 75
 Db 158 ATA-----CCGCCCGGC--GCCCGCGCGGCATCAGAAAGCTGTAGCTGA 200
 QY 76 GlyTrpIleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGlu-IleSe 95
 Db 201 GGC-----GCCCAACAGGATGCGCCCGTAGGATCGGTGTAGC 239
 QY 95 rGlnValValAlaHisGlnLeuGlyThrGlySerValAla-----MetValThrMe 112
 Db 240 GCGGTGCTAGCGCTCTGGAAGTGTAGTCTGTGTTGTGTAACGCGGTTCGCGGTGCA 299
 QY 112 tThrMetArgHisThrAlaGlyGln----- 120
 Db 300 ACAGGTCTTCCACCCCGAGGTCAGCCCGCGCGGTGGCTGAGCGGTAGTCTGCCGGCCA 359
 QY 121 -----ArgLeuHisAspLeuTrpThrGlyLeuSer----- 130
 Db 360 GGTCCACAGAGTTCACAGAACCCGCTTCCCGTGGC-TGGCGCGGTGCTAGTCTGTGTAG 418
 QY 131 -----AlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgTh 144
 Db 419 CGSCTGCTAGCGGTGTGCTCAGGTTGGCTGCGAGCGGTGCGCGGTCCAGTGGCGTTG 478
 QY 144 rGluArgGluMetTyrGly-----CysAspGly 153
 Db 479 GCACGTGCTTCCAGCGGCGGATCAGCCGACGCCCTGGAGTGGCGGAGTGTTCAGG 538
 QY 153 yTyrValArg-----Ala-ValGluIleThrHisGlyLysAsnGlyTrpHisValHisV 171
 Db 539 TACTGGCCGCGGATCTGCTCTGGAAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 580
 QY 171 alHisAlaLeuLeu-----MetPheSerGlyAspValSerGluAsnI 185
 Db 581 TAGTGTGCTTTCAGGTCGAGCCGCACTGCGCGGTGCGGCTGCG-----CGGGAACCA 634
 QY 185 leLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuG 205
 Db 635 TAATCGAGGTCAGGTTCAGCGCC----- 657
 QY 205 lyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGluA 225
 Db 658 -----GCTGCTCTTTACCTTCCGAGGTTGCGCGGTGCGCGGTGCGCGGTGCG 691
 QY 225 laAspGlnValLeu-----AlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetG 243

```

Db      692  GTGACGACGTGATCGATAGACCGCTCGGCTTCGGCAGCATGCGTCCGGGTAGGCTGC 751
QY      243  luValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleA 263
Db      752  GGGTCGGCGAA---GATCGCGCTTCGGGAACTCGCGCAT----- 789
QY      263  laValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheGluPheG 283
Db      790  -----CTGGTTGGGATCTCTGATCCACCAGAGTCCAGGCCGAC-----G 829
QY      283  lySerMetGlyArgAlaIleAlaTrp-----SerArgGlyLeuArgA 298
Db      830  GAAAGGTCGCGAGCGGCTGGTAGACAGCCAGCGTAACGTTGCGCGCGTTTCGGGT 889
QY      298  laArgAlaGlyLeuGlyAla-----GluLeuThrA 308
Db      890  CGCAGGTCGTATGCGCCCGGTGGCGTTGTAGAACTGCTGGCGCAGTCCGCGTTGGCG 949
QY      308  spAlaGlnIleValGluGlnGluGluserAlaProValMetValAlaIleIleProAla 327
Db      950  ATGCGCCTGCTGCTGGCTGGCGCGCGGCACAGCGCGGTCTGCTGTAGTTGGCGCT 1008

```

RESULT 14

```

US-09-335-409-1
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-335-409-1

```

```

Alignment Scores:
Pred. No.:      5.13      Length:      68750
Score:          117.00    Matches:      95
Percent Similarity: 31.25%  Conservative: 45
Best Local Similarity: 21.21% Mismatches:    161
Query Match:      5.84%   Indels:      149
DB:              3       Gaps:       18

```

```

US-10-007-527A-2 (1-379) x US-09-335-409-1 (1-68750)
QY      65  SerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysProCysCysAla--- 83
Db      48236 TCCGGAGGTTCCGGGCGGACCTATGTGGCCAAAGGGTGCCTTCCTCCGCGATGTGCGCAG 48295
QY      83  ----- 83
Db      48296 CTTGGATGCGCGTTCTTCGCCATTTCCCTCTGTGAGCGCATGAGCTGACCGCGMACA 48355
QY      84  -----GlyLysValGlyAlaHisArgAla----- 91
Db      48356 GCGGCTGTTCTGAGGTGAGTGGAGGCGCATCGAGCGGCTGGCCAGGACCCGATGGC 48415
QY      92  -----AspGluIleSerGlnValAlaHis-----GlnLeu 102
Db      48416 GCTGCGGAGAGGCCACCGCGGTGTGTTGGGCGATGATCGGAGCGGAGCACGCCGAGCG 48475
QY      103 GlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeu 122

```

```

Db      48476 GGTGAGGGCCTCGACGACGACGCGCGCTTGTGTACGGCACCCACCGCAACCTGCTCAG 48535
QY      123  HisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrp 142
Db      48536 CGTCCGCGCTGACGCGCTGCTGT-TCCTCTGGGTCTGCACGCGCCGACGATGACCGTGG 48594
QY      143  Arg-----ThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db      48595 ACACGCGCTGCTCGTGTGGTGGGTGCACCTCCCTGCCAGAGCCTCGATGG 48654
QY      161  ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db      48655 GCGAGTGGACACGAGCCCTGCGCGCGGG-TCCAGCGTGTCTTTGTGCGCGCGGTCAATC 48713
QY      181  ValSerGluAsn----- 184
Db      48714 GTGCGGCGTCCGCGATCGGTTTGTTCGCCAGATGGGCGGTGCAAGACGTTCTCGGCC 48773
QY      185  -----IleLeuGluSerPheSer 190
Db      48774 GCTGCAGACGGCTTTGCGCGCGGCGAGGCTGCGCGGTGGTGGTGTCTCAAGCGGCTCCGT 48833
QY      191  AspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeu 210
Db      48834 GACGCGCAGCGCAGCCGC-----GACCCCATCTCTGGCGGTGGTCAGGACGACGCGCATC 48887
QY      211  ArgAsnSerGly-----GlyLeuAspValArgLysIleGlyGlyGluAlaAspGln 227
Db      48888 AACACGATGCGCGCAGCAGCGGCTCACGGTGCCC-----AGCGGTCTGCCACGACAG 48941
QY      228  ValLeu----- 229
Db      48942 CGGTTGCTACGCCAGGCGCTGCGCAAGCGGCGCTGGCGCGCGGAGGTGCAATTCTGTG 49001
QY      230  AlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMetGluVal-----GlySer 246
Db      49002 GAGTGCCACGCGGACGGGACAGCGCTGGTGACCCGATCGAGGTGCAGCGCTGGCGCG 49061
QY      247  GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAla 266
Db      49062 GTGTACGGCGG---GGCGCCCGCGGAGCGCGCTCTGGTGGGCGGTGTCAAGGCC 49118
QY      266  ----- 266
Db      49119 AACCTCGCCACCTGGAGCGCGCGGCTTGGCGCGGTGTCAAGGTGCTCTTGGCG 49178
QY      267  -----ValGlyGlyAspProGlnAlaLeuGluLeu----- 276
Db      49179 CTGGAGCAGCAGCAGATTCGGGCTCAACCGGAGCTCGACGAGCTCAACCCGACATCCCG 49238
QY      277  TrpArgGluPheGluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeu 296
Db      49239 TGGCAGAGCTGCCAGTG-----GCCGTTCGCGAGGCGGTCCCTCTGGCGCGCGCGCG 49295
QY      297  ArgAlaArg-----AlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleVal 312
Db      49296 CGCCCGCTCGTCAGCGCGTGAGCGCTTTCGGCTGAGCGGGACCAACGCGCATGTGGTG 49355
QY      313  GluGlnGluSerAlaProValMetValAlaIleIleProAlaArgSerTrpMetMet 332
Db      49356 TTGGAGAGGCGCGCGGTGGAGCCTGTGGCGCGCGCCCGCGCGCGCGGAGCTG 49415
QY      333  IleArgThrCysAlaProTyrValPheGlyGluIleGlyLeuValGluAlaGlyAla 352
Db      49416 TTCGTCTCTCGCGC-----AAGAGCGCGCGCGCGCTGGATGCGCGAGCA 49460
QY      353  ThrTrpGluAsnLeuArgAspHisLeu-----HisTyrArgLeuProAlaAlaAspVal 370
Db      49461 GCC-----CGGTGCGGAGCCACCTGGAGAGCATGTGAGCTTGGCTTGGCGCGCATGTG 49514
QY      371  ArgProIleIleSerValArg 378

```

Search completed: October 23, 2004, 07:50:26
Job time : 3743 secs

Db 49515 GCGTTCAGCCTGGCGAGCAGCGC 49538

RESULT 15

US-09-568-102-1

; Sequence 1, Application US/09568102

; Patent No. 6346404

; GENERAL INFORMATION:

; APPLICANT: Schupp, Thomas

; APPLICANT: Ligon, James

; APPLICANT: Molnar, Istvan

; APPLICANT: Zirkie, Ross

; APPLICANT: Cyr, Devon

; APPLICANT: Goerlach, Joern

; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

; FILE REFERENCE: 4-30582A

; CURRENT APPLICATION NUMBER: US/09/568,102

; CURRENT FILING DATE: 2000-05-10

; PRIOR APPLICATION NUMBER: 09/335,409

; PRIOR FILING DATE: 1999-06-17

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 68750

; TYPE: DNA

; ORGANISM: Sorangium cellulosum

US-09-568-102-1

Alignment Scores:

Pred. No.:	5.13	Length:	68750
Score:	117.00	Matches:	95
Percent Similarity:	31.25%	Conservative:	45
Best Local Similarity:	21.21%	Mismatches:	161
Query Match:	5.84%	Indels:	149
DB:	3	Gaps:	18

US-10-007-527A-2 (1-379) x US-09-568-102-1 (1-68750)

Qy	65	SerGlyPheGlyLeuArgSerCysGlyysGlyTrrPileCysProCysCysala---	83
Db	48236	TCCGAGGTTCCGGGCCCGACCTATGTGCCAAGGTGCTTCTCCGGCATGTGCGCAG	48295
Qy	83	----	83
Db	48296	CTTGATCGGGGTTCTTCGCCATTTCCTCGTGAGGCGATGAGCTGGACCCGCAACA	48355
Qy	84	-----GlyLysValGlyAlaHisArgAla-----	91
Db	48356	GCGGCTGTTGTCTGAGGTGAGCTGAGTCGGAGCGCATCGAGCGCTGGCCAGGACCCGATGCG	48415
Qy	92	-----AspGluIleSerGlnValValAlaHis-----	102
Db	48416	GCTGCGCAGACGCCACCGGGCGTGTCTGGCATGATCGGAGCGAGCAGCCGCGAGCG	48475
Qy	103	GlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeu	122
Db	48476	GCTGAGGCGCTCGACGACGCGCGGTGTGTACGCGCACACCGGCAACCTGCTCAG	48535
Qy	123	HisAspLeuTrrThrGlyLeuSerAlaAlaTrrLysAlaAlaThrAsnGlyArgArgTrr	142
Db	48536	CGTCGCGCTGGACGGCTGTGCT-TCCTCTGGGCTGTCACGCCCGCAGATACCGGTGG	48594
Qy	143	Arg-----ThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160
Db	48595	ACACGCGCTGCTGCTGCTGCTGGCGTGTGACCTCGCTCCAGAGCGCTCGATTGG	48654
Qy	161	ThrHisGlyLysAsnGlyTrrHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	48655	GCGAGTGCACACGGCCCTGCGCGCGGG-TCCAGCGTGTCTTTTGTGCGCGGTCAATC	48713
Qy	181	ValSerGluAsn-----	184
Db	48714	GTGCGCGCGCTCGCGCATGCGTGTGTTTCGCCAGATGGCGGCTGCAAGAGCTTCTCGGCC	48773

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 05:46:24 ; Search time 607 Seconds
(without alignments)
3199.285 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVAEHLGKDRPPLVLS.....HLVRLPAADVRPPIISVRK 379

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3407233 seqs, 2561960514 residues

Total number of hits satisfying chosen parameters: 6814466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10007527/runat_22102004_114118_26723/app_query.fasta_1.519
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blorum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10007527@cgn1.1.544 @runat_22102004_114118_26723
-NCPU=6 -ICPU=3 -NO_WAPAR -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2005	100.0	1140	14	US-10-007-527A-1	Sequence 1, Appli
2	2005	100.0	1140	14	US-10-007-452-1	Sequence 1, Appli
3	2005	100.0	1140	17	US-10-415-562A-1	Sequence 1, Appli
4	2005	100.0	6334	14	US-10-007-527A-5	Sequence 5, Appli
5	2005	100.0	6334	14	US-10-007-452-5	Sequence 5, Appli
6	2005	100.0	6334	17	US-10-415-562A-5	Sequence 5, Appli
7	2005	100.0	9652	14	US-10-007-527A-7	Sequence 7, Appli
8	2005	100.0	9652	14	US-10-007-452-7	Sequence 7, Appli
9	2005	100.0	9652	17	US-10-415-562A-7	Sequence 7, Appli
10	2005	100.0	11241	14	US-10-007-527A-6	Sequence 6, Appli
11	2005	100.0	11241	14	US-10-007-452-6	Sequence 6, Appli
12	2005	100.0	11241	17	US-10-415-562A-6	Sequence 6, Appli
13	337	16.8	7560	9	US-09-754-112A-2	Sequence 2, Appli
14	337	16.8	12738	9	US-09-754-112A-1	Sequence 1, Appli
15	326	16.3	4447	9	US-09-835-381-5	Sequence 5, Appli
16	326	16.3	4447	15	US-10-196-232-5	Sequence 5, Appli
17	318	15.9	4447	9	US-09-835-381-7	Sequence 7, Appli
18	318	15.9	4447	15	US-10-196-232-6	Sequence 6, Appli
19	296	14.8	8500	16	US-10-296-947-7	Sequence 7, Appli
20	282	14.1	2401	9	US-09-826-191-6	Sequence 6, Appli
21	282	14.1	2401	14	US-10-263-666-6	Sequence 6, Appli
22	282	14.1	2401	14	US-10-261-481-6	Sequence 6, Appli
23	282	14.1	2401	14	US-10-261-942-6	Sequence 6, Appli
24	282	14.1	4005	9	US-09-826-191-2	Sequence 2, Appli
25	282	14.1	4005	14	US-10-263-666-2	Sequence 2, Appli
26	282	14.1	4005	14	US-10-261-481-2	Sequence 2, Appli
27	282	14.1	4005	14	US-10-261-942-2	Sequence 2, Appli
28	138	6.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
29	136	6.8	4368	17	US-10-437-963-33355	Sequence 33355, A
30	133	6.6	2256646	17	US-10-470-565-1	Sequence 96272, A
31	128.5	6.4	1763	17	US-10-437-963-96272	Sequence 1, Appli
32	125	6.2	2256646	17	US-10-470-565-1	Sequence 1, Appli
33	123	6.1	1379	17	US-10-437-963-95384	Sequence 95384, A
34	123	6.1	25582	17	US-10-322-281-578	Sequence 578, App
35	122.5	6.1	1540	17	US-10-437-963-32408	Sequence 32408, A
36	122	6.1	2871	9	US-09-815-242-7827	Sequence 7827, Ap
37	122	6.1	2871	16	US-10-282-122A-30310	Sequence 30310, A
38	119	5.9	2427	16	US-10-282-122A-23453	Sequence 23453, A
39	118.5	5.9	1386	17	US-10-437-963-48219	Sequence 48219, A
40	118	5.9	1386	15	US-10-314-657-1	Sequence 1, Appli
41	117	5.8	68750	13	US-10-014-717-1	Sequence 7, Appli
42	116	5.8	1542	9	US-09-833-745-7	Sequence 12, Appli
43	116	5.8	15738	15	US-10-329-079-12	Sequence 6, Appli
44	116	5.8	37360	15	US-10-329-079-6	Sequence 53737, A
45	115.5	5.8	1661	17	US-10-437-963-53737	

ALIGNMENTS

RESULT 1
US-10-007-527A-1
; Sequence 1, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Rhodococcus AN12

US-10-007-527A-1									
Alignment Scores:									
Pred. No.:	8.33e-216	Length:	1140						
Score:	2005.00	Matches:	379						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	14	Gaps:	0						
US-10-007-527A-2 (1-379) x US-10-007-527A-1 (1-1140)									
QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer	20						
DB	1	ATGACCAAGCGTAAGTGTGAACACCTTTCGGCAAGAACCGCGCTCCGCTCGTGTGCG	60						
QY	21	SerAspLysArgGlyLleArgHisGluLeuArgProLysLeuGlnIleThrThrSer	40						
DB	61	TCCGATAAGCGCGCATCCGCGACGAACTTTCGGACCCCAACTTCAACAAATCACCACGTCA	120						
QY	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60						
DB	121	GAACATTATACGCTGTGGCGCGCATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC	180						
QY	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro	80						
DB	181	GGTCCGAAAGGTTCTGGATTCCGAGGCTTCTGTTCTCGGAAAGGGCTGGATCTGCC	240						
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100						
DB	241	TGCTGTGGGAAAGTCCGTCGATCGTGCACATCGTCGACAGGAAATTTCTCAAGTTGTGCTCAT	300						
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120						
DB	301	CAACTCGGACGTGGATCTGTTGCGATGGTACGATGACCATCGCCATACAGCTGGTCAG	360						
QY	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140						
DB	361	CGGCTCCACGACCTATGACTGGACTTTTCGGCAGGCTTGAAAGCTGCGACCAACGGTGTG	420						
QY	141	ArgTrpArgThrGluArgGluMetTrpGlyCysAspGlyTrpValArgAlaValGluIle	160						
DB	421	CGTGGCGTACGGAACGTAATGTACGCTGCGACGATAGTGGCGCTGTGAAATC	480						
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180						
DB	481	ACTCAGGAAACACGGCTGGCAGCTCCACGTTCCAGCGCTACTCATGTTTCAGTGTGAC	540						
QY	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200						
DB	541	GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGGATGTCGATCGGTGGACTTCCAA	600						
QY	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220						
DB	601	CTCGTATCTCGGATTGCTGGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG	660						
QY	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyVal	240						
DB	661	ATCGCGGTGAAAGCTCATCAAGTTCTCGTGGCTATCTCAGCAAAATTTGCATCTGGCGTT	720						
QY	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260						
DB	721	GGTATGAGGTTGGTAGTGGCGAGAAAGTGGTCGACATGGCAACCGTGCACCCTGG	780						
QY	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280						
DB	781	GAATCGCTGTTGATGAGTGGCGGGATCCACACGGTGGAACTGTGGCGAGAATTT	840						
QY	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300						
DB	841	GAGTTTGGTTCGATGGGACGTCCGGCAATCGCGTGGTCCCGTGGATTGGCTGCCGAGCT	900						
QY	301	GlyLeuGlyAlaGluLeuThrArgAlaGlnIleValGlnGlnGluSerAlaProVal	320						

US-10-007-527A-1									
Alignment Scores:									
Pred. No.:	8.33e-216	Length:	1140						
Score:	2005.00	Matches:	379						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	14	Gaps:	0						
US-10-007-527A-2 (1-379) x US-10-007-527A-1 (1-1140)									
QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer	20						
DB	1	ATGACCAAGCGTAAGTGTGAACACCTTTCGGCAAGAACCGCGCTCCGCTCGTGTGCG	60						
QY	21	SerAspLysArgGlyLleArgHisGluLeuArgProLysLeuGlnIleThrThrSer	40						
DB	61	TCCGATAAGCGCGCATCCGCGACGAACTTTCGGACCCCAACTTCAACAAATCACCACGTCA	120						
QY	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60						
DB	121	GAACATTATACGCTGTGGCGCGCATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC	180						
QY	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro	80						
DB	181	GGTCCGAAAGGTTCTGGATTCCGAGGCTTCTGTTCTCGGAAAGGGCTGGATCTGCC	240						
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100						
DB	241	TGCTGTGGGAAAGTCCGTCGATCGTGCACATCGTCGACAGGAAATTTCTCAAGTTGTGCTCAT	300						
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120						
DB	301	CAACTCGGACGTGGATCTGTTGCGATGGTACGATGACCATCGCCATACAGCTGGTCAG	360						
QY	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140						
DB	361	CGGCTCCACGACCTATGACTGGACTTTTCGGCAGGCTTGAAAGCTGCGACCAACGGTGTG	420						
QY	141	ArgTrpArgThrGluArgGluMetTrpGlyCysAspGlyTrpValArgAlaValGluIle	160						
DB	421	CGTGGCGTACGGAACGTAATGTACGCTGCGACGATAGTGGCGCTGTGAAATC	480						
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180						
DB	481	ACTCAGGAAACACGGCTGGCAGCTCCACGTTCCAGCGCTACTCATGTTTCAGTGTGAC	540						
QY	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200						
DB	541	GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGGATGTCGATCGGTGGACTTCCAA	600						
QY	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220						
DB	601	CTCGTATCTCGGATTGCTGGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG	660						
QY	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyVal	240						
DB	661	ATCGCGGTGAAAGCTCATCAAGTTCTCGTGGCTATCTCAGCAAAATTTGCATCTGGCGTT	720						
QY	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260						
DB	721	GGTATGAGGTTGGTAGTGGCGAGAAAGTGGTCGACATGGCAACCGTGCACCCTGG	780						
QY	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280						
DB	781	GAATCGCTGTTGATGAGTGGCGGGATCCACACGGTGGAACTGTGGCGAGAATTT	840						
QY	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300						
DB	841	GAGTTTGGTTCGATGGGACGTCCGGCAATCGCGTGGTCCCGTGGATTGGCTGCCGAGCT	900						
QY	301	GlyLeuGlyAlaGluLeuThrArgAlaGlnIleValGlnGlnGluSerAlaProVal	320						

US-10-007-527A-1									
Alignment Scores:									
Pred. No.:	8.33e-216	Length:	1140						
Score:	2005.00	Matches:	379						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	14	Gaps:	0						
US-10-007-527A-2 (1-379) x US-10-007-527A-1 (1-1140)									
QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer	20						
DB	1	ATGACCAAGCGTAAGTGTGAACACCTTTCGGCAAGAACCGCGCTCCGCTCGTGTGCG	60						
QY	21	SerAspLysArgGlyLleArgHisGluLeuArgProLysLeuGlnIleThrThrSer	40						
DB	61	TCCGATAAGCGCGCATCCGCGACGAACTTTCGGACCCCAACTTCAACAAATCACCACGTCA	120						
QY	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60						
DB	121	GAACATTATACGCTGTGGCGCGCATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC	180						
QY	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro	80						
DB	181	GGTCCGAAAGGTTCTGGATTCCGAGGCTTCTGTTCTCGGAAAGGGCTGGATCTGCC	240						
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100						
DB	241	TGCTGTGGGAAAGTCCGTCGATCGTGCACATCGTCGACAGGAAATTTCTCAAGTTGTGCTCAT	300						
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120						
DB	301	CAACTCGGACGTGGATCTGTTGCGATGGTACGATGACCATCGCCATACAGCTGGTCAG	360						
QY	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140						
DB	361	CGGCTCCACGACCTATGACTGGACTTTTCGGCAGGCTTGAAAGCTGCGACCAACGGTGTG	420						
QY	141	ArgTrpArgThrGluArgGluMetTrpGlyCysAspGlyTrpValArgAlaValGluIle	160						
DB	421	CGTGGCGTACGGAACGTAATGTACGCTGCGACGATAGTGGCGCTGTGAAATC	480						
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180						
DB	481	ACTCAGGAAACACGGCTGGCAGCTCCACGTTCCAGCGCTACTCATGTTTCAGTGTGAC	540						
QY	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200						
DB	541	GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGGATGTCGATCGGTGGACTTCCAA	600						
QY	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220						
DB	601	CTCGTATCTCGGATTGCTGGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG	660						
QY	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyVal	240						
DB	661	ATCGCGGTGAAAGCTCATCAAGTTCTCGTGGCTATCTCAGCAAAATTTGCATCTGGCGTT	720						
QY	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260						
DB	721	GGTATGAGGTTGGTAGTGGCGAGAAAGTGGTCGACATGGCAACCGTGCACCCTGG	780						
QY	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280						
DB	781	GAATCGCTGTTGATGAGTGGCGGGATCCACACGGTGGAACTGTGGCGAGAATTT	840						
QY	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300						
DB	841	GAGTTTGGTTCGATGGGACGTCCGGCAATCGCGTGGTCCCGTGGATTGGCTGCCGAGCT	900						
QY	301	GlyLeuGlyAlaGluLeuThrArgAlaGlnIleValGlnGlnGluSerAlaProVal	320						

US-10-007-527A-1									
Alignment Scores:									
Pred. No.:	8.33e-216	Length:	1140						
Score:	2005.00	Matches:	379						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	14	Gaps:	0						
US-10-007-527A-2 (1-379) x US-10-007-527A-1 (1-1140)									
QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer	20						
DB	1	ATGACCAAGCGTAAGTGTGAACACCTTTCGGCAAGAACCGCGCTCCGCTCGTGTGCG	60						
QY	21	SerAspLysArgGlyLleArgHisGluLeuArgProLysLeuGlnIleThrThrSer	40						
DB	61	TCCGATAAGCGCGCATCCGCGACGAACTTTCGGACCCCAACTTCAACAAATCACCACGTCA	120						
QY	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60						
DB	121	GAACATTATACGCTGTGGCGCGCATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC	180						
QY	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro	80						
DB	181	GGTCCGAAAGGTTCTGGATTCCGAGGCTTCTGTTCTCGGAAAGGGCTGGATCTGCC	240						
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100						
DB	241	TGCTGTGGGAAAGTCCGTCGATCGTGCACATCGTCGACAGGAAATTTCTCAAGTTGTGCTCAT	300						
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120						
DB	301	CAACTCGGACGTGGATCTGTTGCGATGGTACGATGACCATCGCCATACAGCTGGTCAG	360						
QY	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140						
DB	361	CGGCTCCACGACCTATGACTGGACTTTTCGGCAGGCTTGAAAGCTGCGACCAACGGTGTG	420						
QY	141	ArgTrpArgThrGluArgGluMetTrpGlyCysAspGlyTrpValArgAlaValGluIle	160						
DB	421	CGTGGCGTACGGAACGTAATGTACGCTGCGACGATAGTGGCGCTGTGAAATC	480						
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180						
DB	481	ACTCAGGAAACACGGCTGGCAGCTCCACGTTCCAGCGCTACTCATGTTTCAGTGTGAC	540						
QY	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200						
DB	541	GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGGATGTCGATCGGTGGACTTCCAA	600						
QY	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220						
DB	601	CTCGTATCTCGGATTGCTGGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG	660						
QY	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyVal	240						
DB	661	ATCGCGGTGAAAGCTCATCAAGTTCTCGTGGCTATCTCAGCAAAATTTGCATCTGGCGTT	720						
QY	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260						
DB	721	GGTATGAGGTTGGTAGTGGCGAGAAAGTGGTCGACATGGCAACCGTGCACCCTGG	780						
QY	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280						
DB	781	GAATCGCTGTTGATGAGTGGCGGGATCCACACGGTGGAACTGTGGCGAGAATTT	840						
QY	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300						
DB	841	GAGTTTGGTTCGATGGGACGTCCGGCAATCGCGTGGTCCCGTGGATTGGCTGCCGAGCT	900						
QY	301	GlyLeuGlyAlaGluLeuThrArgAlaGlnIleValGlnGlnGluSerAlaProVal	320						

US-10-007-527A-1									
Alignment Scores:									
Pred. No.:	8.33e-216	Length:	1140						
Score:	2005.00	Matches:	379						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	14	Gaps:	0						
US-10-007-527A-2 (1-379) x US-10-007-527A-1 (1-1140)									
QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer	20						
DB	1	ATGACCAAGCGTAAGTGTGAACACCTTTCGGCAAGAACCGCGCTCCGCTCGTGTGCG	60						
QY	21	SerAspLysArgGlyLleArgHisGluLeuArgProLysLeuGlnIleThrThrSer	40						
DB	61	TCCGATAAGCGCGCATCCGCGACGAACTTTCGGACCCCAACTTCAACAAATCACCACGTCA	120						
QY	41	GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn	60						
DB	121	GAACATTATACGCTGTGGCGCGCATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC	180						
QY	61	GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro	80						
DB	181	GGTCCGAAAGGTTCTGGATTCCGAGGCTTCTGTTCTCGGAAAGGGCTGGATCTGCC	240						
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis	100						
DB	241	TGCTGTGGGAAAGTCCGTCGATCGTGCACATCGTCGACAGGAAATTTCTCAAGTTGTGCTCAT	300						
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120						
DB	301	CAACTCGGACGTGGATCTGTTGCGATGGTACGATGACCATCGCCATACAGCTGGTCAG	360						
QY	121	ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg	140						
DB	361	CGGCTCCACGACCTATGACTGGACTTTTCGGCAGGCTTGAAAGCTGCGACCAACGGTGTG	420						
QY	141	ArgTrpArgThrGluArgGluMetTrpGlyCysAspGlyTrpValArgAlaValGluIle	160						
DB	421	CGTGGCGTACGGAACGTAATGTACGCTGCGACGATAGTGGCGCTGTGAAATC	480						
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180						
DB	481	ACTCAGGAAACACGGCTGGCAGCTCCACGTTCCAGCGCTACTCATGTTTCAGTGTGAC	540						
QY	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200						
DB	541	GTGAGTGAGAACATCTCGAATCTTCTCGGATGCGGATGTCGATCGGTGGACTTCCAA	600						
QY	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys	220						
DB	601	CTCGTATCTCGGATTGCTGGCGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG	660						
QY	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTrpLeuThrLysIleAlaSerGlyVal	240						
DB	661	ATCGCGGTGAAAGCTCATCAAGTTCTCGTGGCTATCTCAGCAAAATTTGCATCTGGCGTT	720						
QY	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260						
DB	721	GGTATGAGGTTGGTAGTGGCGAGAAAGTGGTCGACATGGCAACCGTGCACCCTGG	780						
QY	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280						
DB	781	GAATCGCTGTTGATGAGTGGCGGGATCCACACGGTGGAACTGTGGCGAGAATTT	840						
QY	281	GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgG							

Db 361 CGGCTCCAGACCTATGGACTGGACTTTCGGAGCCTGGAAAGCTGGCAACACGGTCGT 420
 QY 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 Db 421 CGTGGCGTACGGAACTGAAATGTACGGCTGGACGGATACGTGGCGCTGTGAATC 480
 QY 161 ThrHisGlyLysAsnGlyTyrHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 Db 481 ACTCAGGAAAAACGGCTGGCAGCTCCAGTTTCACGGCTACTCATGTTTCAGTGGTAC 540
 QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
 Db 541 GTGAGTGAGAACATCCTCGAATCTTCTCGATGCGATGTTCGATCGTGGACTTCCAA 600
 QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLysLeuAspValArgLys 220
 Db 601 CTCGTATCTCTGGATTGCTGGCCACCTACGTAATTCGGGTGGTCTCGATGTACGAAAG 660
 QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
 Db 661 ATCGCGGTGAAGCTCATCAAGTTCTCGCTGGTATCTGACGAAATTCGATCTGGCGTT 720
 QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
 Db 721 GGTATGGAGTTGGTAGTGGCGACGGAAGTGTGCGACATGGCAACCGTGCCACCTCG 780
 QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
 Db 781 GAAATCGCTGTTGATGCACTGGCGGGATCCACAAAGCTTGGAACTGTGGCGAGAATTT 840
 QY 281 GluPheGlySerMetGlyArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla 300
 Db 841 GAGTTTGGTTCGATGGAGCTCGGCAATCGCGTCCGTGGTCCGTGGATGGTCCCGAGCT 900
 QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
 Db 901 GGTCTTGGGCGAAGACTAACAGATGCTCAGATCGTTGAGCAGGAAGAATCTGCGCTTACGTC 960
 QY 321 MetValAlaIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal 340
 Db 961 ATGGTTGCGATCATTCGCGCGGATCGTGGATGATGATTCGGACTTGTGCGCTTACGTC 1020
 QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
 Db 1021 TTCGCGGAGATCCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAAATCTTCGTGATCAC 1080
 QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
 Db 1081 TTGCATTATCATTGCCGCGAGCGATGTCGGCGCCCGGATAATATATCGTTTCGCAAG 1137

RESULT 3

US-10-415-562A-1

; Sequence 1, Application US/10415562A

; Publication No. US20040115661A1

; GENERAL INFORMATION:

; APPLICANT: E.I. du Pont De Nemours and Company

; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

; FILE REFERENCE: CL1709 US PCT

; CURRENT APPLICATION NUMBER: US/10/415,562A

; CURRENT FILING DATE: 2003-11-17

; PRIOR APPLICATION NUMBER: 60/254,868

; PRIOR FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1140

; TYPE: DNA

; ORGANISM: Rhodococcus AN12

US-10-415-562A-1

Alignment Scores:

Pred. No.: 8.33e-216 Length: 1140

Score: 2005.00 Matches: 379
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-415-562A-1 (1-1140)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
 Db 1 ATGACACGCGTAACTAGTCTGACACACCTTTCGGGCAAGACCGGCTCCCGTCCTCGTGTGC 60
 QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
 Db 61 TCGGATAAGCGCGCATCCGGCACGAACCTCGCACCCAACTTCAACAAATCACACGTCAC 120
 QY 41 GluThrPheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsn 60
 Db 121 GAAACATTTAAACCCCTGTGGCCGGCGATTTCCTGCGAAAGGGCTGGATCTGCCCC 180
 QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
 Db 181 GGTCCGAAAGGTTCTGGATTCCGAGGCCTTCGTTCTGCGAAAGGGCTGGATCTGCCCC 240
 QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
 Db 241 TGCTGTGCGGAAAAAGTTCGGTTCGAGGCCTTCGTTCTGCGAAAGGGCTGGATCTGCCCC 300
 QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
 Db 301 CAACTCCGGAGCTGGATCTGTTGCGATGGTGCATGACCATGCGCATACAGCTGGTCAG 360
 QY 121 ArgLeuHisAspLeuTyrThrGlyLeuSerAlaAlaTyrLysAlaAlaThrAsnGlyArg 140
 Db 361 CGGCTCCAGCCTATGGACTTGGACTTTCGGCAGGCTGGAAAGCTTGGACCAACGGTGGT 420
 QY 141 ArgTyrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
 Db 421 CGTTGGCGTACGGAACGTGAAATGTACGGCTCGCAGCGATACGTGCGCGCTGCTGAAATC 480
 QY 161 ThrHisGlyLysAsnGlyTyrHisValHisAlaLeuLeuMetPheSerGlyAsp 180
 Db 481 ACTCAGGAAAAACGGCTGGCAGCTCCACGTTTCACGCGCTACTCATGTTTCAGTGGTGC 540
 QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
 Db 541 GTGAGTGAGACATCCTCGAATCTTTCGGATGCGATGTTTCGATCGGTGGACTTCCAAA 600
 QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
 Db 601 CTCGTATCTCTGGGATTGCTGGCCCACTACGTAAATTCGGGTGGTCTCGATGTACGAAAG 660
 QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
 Db 661 ATCGCGCGTGAAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTTGCACTTGGCGTT 720
 QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
 Db 721 GGTATGAGGTTGGTAGTGGCGCGGAAAAAGTGGTGCATGCGAACCCGTGGACCCCTGG 780
 QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
 Db 781 GAAATCGCTGTTGATGAGTGGCGGGATCCCAACAGCTTGGAACTGTGCGCGAGAATTT 840
 QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTyrSerArgGlyLeuArgAlaArgAla 300
 Db 841 GAGTTTGGTTCGATGGAGCGCTGGGCAATCCGCTGGTTCGCTGGATGGCTGCCCGAGCT 900
 QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluLeuGluSerAlaProVal 320
 Db 901 GGTCTTGGGCGCAGAACTAACAGATGCTCAGATCGTTGAGCAGGAAGAATCTGCCCGCGTC 960
 QY 321 MetValAlaIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal 340

Db 961 ATGTTTCGATCAATTCGGCGCGATCGTGATGATGATTCGGATTTGTGGCCCTTACGTC 1020
Qy 341 PheGlyGluLeuLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 1021 TTCGGCGAGATCCTCGGACTCGTCGAGCTGGCGGACTTGGGAAATCTTCGTGATCAC 1080
Qy 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProLeuLeuSerValArgLys 379
Db 1081 TTGCATTATCGATTGCCCGCAGCGAGTGTGGCGGCCCGGATAATATCGGTTTCGCAAG 1137
RESULT 4
US-10-007-527A-5/c
; Sequence 5, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-007-527A-5
Alignment Scores:
Pred. No.: 7,46e-215 Length: 6334
Score: 2005.00 Matches: 379
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 14
US-10-007-527A-2 (1-379) x US-10-007-527A-5 (1-6334)
Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
Db 3051 ATGACCAGCGTAACTGCTGAACACCTTTCGGCAAGACCGGCTCCCGTCTCGTGTCG 2992
Qy 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnThrThrSer 40
Db 2991 TCCGATAAGCGCGCATCCGGCAGAACTGCGACCCCAAACTTCAACAAATCACACGTCA 2932
Qy 41 GluThrPheAsnAlaCysGlyArgProLysLeuValAsnGlyValThrLeuValAsn 60
Db 2931 GAAACATTAAACCTGTGCGCGCGGATTTCTGGCGTGAACGGTGTGACCAATGTCAAC 2872
Qy 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyThrLysCysPro 80
Db 2871 GGTCGAAAGGTTCTCGATTTCGGAGGCTTCGTTCTTCCTGCGGAAAGGGCTGATCTGCC 2812
Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLeuSerGlnValAlaHis 100
Db 2811 TGCTGTGCGGAAAGTCGGTGCACATCGTGACAGCAAAATTTCTCAAGTTGTGCTCAT 2752
Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 2751 CAATCGGGACTGATCTGTTCGATGATGATGATGATGATGATGATGATGATGATGATG 2692
Qy 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140
Db 2691 CGGCTCCACGACCTATGACTGAGCTTTTCGGCAGCTGGAAGCTGCCACCAACGGTCTG 2632
Qy 141 ArgTrpArgGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluLeu 160

Db 2631 CGTTGGCGTACGGAACCGTAAATGTACGGCTGCGACGATACGTCGCGCTGTTGAAATC 2572
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 2571 ACTCACGGAATAACCGCTGGCAGCTCCAGCTTCCCGGCTACTCATGTTTCAGTGGTGAC 2512
Qy 181 ValSerGluAsnLeuLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db 2511 GTGAGTGAGAACATCTCTCGAATCTTCTCGATCGATGTTTCGATCGGTGACTTCCAAA 2452
Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 2451 CTCGTATCTCTGGGATTTGCTGCGCCACTAGCTAATTCGGGTGGTCTCGATGACGAAG 2392
Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 2391 ATCGGCGGTGAAGCTGATCAAGTTCTCGCTCGATATCTGACGAAATTCATCTGCGCT 2332
Qy 241 GlyMetGluValGlySerGlyAspLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 2331 GGTATGGAGGTTGGTAGTGGCGACGGAATAAGTGTGACATGCGAACCGTGCACCTGG 2272
Qy 261 GluIleAlaValAspAlaValGlyClyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 2271 GAAATCGCTGTTGATGTCAGTGGCGGGATCCACAGCGTTGGAACCTGTGGCGAAATTT 2212
Qy 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 2211 GAGTTTGGTTCGATGGGACGTCGGCAATCGGCTGGTCCCCTGGATTCGTCGCCGAGCT 2152
Qy 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 2151 GGTCTTGGGCGAAGAACTAACAGATGCTCAGATCGTTGACGAGAAATCTGCCCGGTC 2092
Qy 321 MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTrpVal 340
Db 2091 ATGGTTTCGATCATTCGGCGCGATCGTGGATGATGATTCGGACTTGTGGCCCTTACGIC 2032
Qy 341 PheGlyGluLeuLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 2031 TTCGGCGAGATCTCGGACTCGTCGAGCTCGGAGCTGGGAAATCTTCGTGATCAC 1972
Qy 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProLysSerValArgLys 379
Db 1971 TTGCATTATCGATTGCCCGCAGCGATGTCGGCCCCCGGATATATATCGTTCGCAAG 1915
RESULT 5
US-10-007-452-5/c
; Sequence 5, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 6334
; TYPE: DNA
; ORGANISM: Rhodococcus AN12
US-10-007-452-5
Alignment Scores:
Pred. No.: 7,46e-215 Length: 6334
Score: 2005.00 Matches: 379

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-10-007-527A-2 (1-379) x US-10-007-452-5 (1-6334)			
QY	1	MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer	20
Db	3051	ATGACGAGCGTAAGTGTGAACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTG	2992
QY	21	SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGlnThrThrSer	40
Db	2991	TCCGATAAGCGCGCATCCGCGACGAACCTCGACCCCAAACTTCAACAAATCACCGTCA	2932
QY	41	GluThrPheAsnAlaCysGlyArgProLysLeuValAsnGlyValThrLeuValAsn	60
Db	2931	GAACAATTTAACCCCTGTGCGCGCGGATTTCTGGCGTGAACCGGTGACCATTTGTCAAC	2872
QY	61	GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpLysCysPro	80
Db	2871	GGTCCGAAAGGTTCTCGATTCCGAGGCTTGGTCTCTGGGAAAGGCTGGATCTGCCCC	2812
QY	81	CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValValAlaHis	100
Db	2811	TGCTGTGCGGAAAGTGTGCGTGCATCGTCAGACGAAATTTCTCAAGTTGTGTCTCAT	2752
QY	101	GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln	120
Db	2751	CAACTCGGACGCTGGATCTGTTGCGATGGTGAACATCGCATCGCCATACAGCTGTCTAG	2692
QY	121	ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg	140
Db	2691	CGGCTCCACGACCATGTGACCTGGACTTTCGGCAGCCTGGAAGCTGCGACCAACCGTCT	2632
QY	141	ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle	160
Db	2631	CGTTGGCGTACGGAACGTGAATGTACGGCTGCGAGGATACGTGCGCGCTGTGGAATC	2572
QY	161	ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp	180
Db	2571	ACTCAGGAAAGAAACCGCTGGCAGCTCCACGTTCCACGCTACTCATGTTAGTGGTAC	2512
QY	181	ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys	200
Db	2511	GTGAGTGAGAACATCTCGAATCTTCTCGATGCGATGTTGATCGGTGGACTTCCAAA	2452
QY	201	LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyLeuAspValArgLys	220
Db	2451	CTCGTATCTCTGGGATTGTCGCGCACTACGTAATTCGGGTGTTCTCGATGTACGAAAG	2392
QY	221	IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal	240
Db	2391	ATCGCGGCTGAGCTGATCAAGTTCTCGCTCGCTATCTGACGAAATTTGCACTCGCGTT	2332
QY	241	GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp	260
Db	2331	GGTATGAGGTTGTTAGTGGCGACGGAAGTGTGACATGCAACCGTGCACCTCGG	2272
QY	261	GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe	280
Db	2271	GAATTCGCTGTTATGTCAGTGGCGGGATCCCAACGCTTGAACCTGGCGGAGATTT	2212
QY	281	GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla	300
Db	2211	GAGTTTGGTTCGATGGGACGTGGGCAATCGCTGGTCCCGTGGATTTGCGTGGCCGAGCT	2152
QY	301	GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluLeuSerAlaProVal	320
Db	2151	GGTCTTGGGCGAGAACTAAACAGATGCTCAGATGTTGAGCAGGAAGAAATCTGCCCGGTC	2092
QY	321	MetValAlaIleIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal	340
Db			

Db 2091 ATGTTTCCGATCATTCGCGCGCATCGTGGATGATGATTCGGAATTGTGGCCCTTACGTC 2032

QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluLeuLeuArgAspHis 360

Db 2031 TTCGGCGAGATCTCGGACTCTCGAAGCTGGCGGACCTTGGGAAATCTTCGTGATCAC 1972

QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleLeuSerValArgLys 379

Db 1971 TTTCATTATCGATTGCGCGCATGTGCGGCCCGCATATATATATCGTTGCAAG 1915

RESULT 6

US-10-415-562A-5/c

; Sequence 5, Application US/10415562A

; Publication No. US20040115661A1

; GENERAL INFORMATION:

; APPLICANT: E.I. du Pont De Nemours and Company

; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors

; FILE REFERENCE: CL1709 US PCT

; CURRENT APPLICATION NUMBER: US/10/415,562A

; PRIOR FILING DATE: 2003-11-17

; PRIOR APPLICATION NUMBER: 60/254,868

; FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 5

; LENGTH: 6334

; TYPE: DNA

; ORGANISM: Rhodococcus AN12

US-10-415-562A-5

Alignment Scores:

Pred. No.: 7,46e-215 Length: 6334

Score: 2005.00 Matches: 379

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 17 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-415-562A-5 (1-6334)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20

Db 3051 ATGACGAGCGTAAGTGTGAACACCTTTCGGCAAGACCGGCTCCCGTCTCGTGTG 2992

QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnThrThrSer 40

Db 2991 TCCGATAAGCGCGCATCCGCGACGAACCTTTCGGGCAAGACCGGCTCCCGTCTCGTGTG 2932

QY 41 GluThrPheAsnAlaCysGlyArgProLysLeuValAsnGlyValThrIleValAsn 60

Db 2931 GAAACATTTAACCGCTTGGCGCGCGATTTCTGGCGTGAACCGGTGACCATTTGTCAAC 2872

QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpLysCysPro 80

Db 2871 GGTCGAAAGGTTCTCGATTCCGAGGCTTGGTCTCTGGGAAAGGCTGGATCTGCCCC 2812

QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValValAlaHis 100

Db 2811 TGCTGTGCGGAAAGTGTGCGTGCATCGTCAGACGAAATTTCTCAAGTTGTGTCTCAT 2752

QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120

Db 2751 CAACTCGGACGCTGGATCTGTTGCGATGGTGAACATCGCATCGCCATACAGCTGTCTAG 2692

QY 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140

Db 2691 CGGCTCCACGACCATGTGACCTGGACTTTCGGCAGCCTGGAAGCTGCGACCAACCGTCT 2632

QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160

Db 2631 CGTTGGCGTACGGAACGTGAATGTACGGCTGCGAGGATACGTGCGCGCTGTGGAATC 2572

QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180

Db 4032 TTCGGCGAGATCCTCGGACTCTGCGAGCTGGCGGCGACTTGGGAAAATCTTCGTGATCAC 3973
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProPheSerValArgLys 379
Db 3972 TTGCATTATCGATTGCCCGAGCGATGTCGGCGCCCGGATAATATATCGGTTCGCAAG 3916
RESULT 8
US-10-007-452-7/c
; Sequence 7, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-452-7
Alignment Scores:
Pred. No.: 1,28e-214 Length: 9652
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-007-527A-2 (1-379) x US-10-007-452-7 (1-9652)
QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
Db 5052 ATGACCGGTAAGTGTGACACCTTCCGGCAAGACCGGCTCCCGTCTCGTGTG 4993
QY 21 SerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGlnGluThrSer 40
Db 4992 TCCGATAGCGGCATCGGCACGAATCGGACCAACTTCAACAATCACCACGTCA 4933
QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 4932 GAAACATTAAAGCTGTGGCGGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTCAC 4873
QY 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyTyrPileCysPro 80
Db 4872 GTCCGGAAGGTTCTGGAATCGGAGGCTTCTGCTCCGGAAAGGGCTGATCTGCCCC 4813
QY 81 CysCysAlaGlyLysValGlyValAlaHisArgAlaAspGluLeuSerGlnValAlaHis 100
Db 4812 TGCTGTGGGAAAGTCGGTGCATCTGTCACAGCAAAATTTCTCAAGTTGTTCAT 4753
QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 4752 CAATCGGGATGGATCTGTTCGATGGTACGATGACCATGCCCATACAGCTGGTTCAG 4693
QY 121 ArgLeuHisAspLeuTyrThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 4692 CGGCTCCACACCTATGAGCTGGACTTTCGACGCTGGAAAGCTGGCAACCGTCTG 4633
QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluLe 160
Db 4632 CGTTGGCGTACGNAAGTGAATGTACGGCTGCGACGGATACGTGCGCTGTGTAATC 4573
QY 161 ThrHisGlyLysAsnGlyTyrHisValHisValHisValHisValHisValHisValHis 180

Db 4572 ACTCAGGAAAAAAGCGCTGGCACGCTCCAGCTTCCAGCGCTACTCATGTTTCAGTGTGAC 4513
QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTyrThrSerLys 200
Db 4512 GTGAGTGAGAACATCTCGATCTCTTCGATGCGATGTTTCGATCGGTGGACTTCCAAA 4453
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 4452 CTCGTATCTCTGGGATTTCTGCGCCACTACGTAATTCGGGTGCTCTCGATGTACGAAG 4393
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 4392 ATCGCGGTGAAGCTCATCAAGTTCTCGCTGCGTATCTCACGAAAAATTCGATCTGCGTT 4333
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyr 260
Db 4332 GGTATGGAGTTGTTAGTGGCGACGGAAGTGTGACATGGCAACCGTGACCTGG 4273
QY 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTyrArgGluPhe 280
Db 4272 GAAATCGCTTTGATGACGTGGCGGATCCACAAGCGTTGGAACCTGTGGCGAGATTT 4213
QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 4212 GAGTTTGGTTTCATGGGACGTGGGCAATCGCGTGTGCGTGGATTTGCGTCCGAGCT 4153
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 4152 GGTCTTGGCGAGAACTAACAGATCTCAGATCTGTGAGCAGAAAGATCTGCCCGGTC 4093
QY 321 MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTyrVal 340
Db 4092 ATGTTGCGATCATTCGCGCGCATCGTGGATGATGATTTCGACCTTGTGGCTTACGTC 4033
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 4032 TTCGCGAGATCTCTCGGACTCGTGAAGCTGGCGGACTTGGGAAAATCTTCGTGATCAC 3973
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleIleSerValArgLys 379
Db 3972 TTGCATTATCGATTGCCCGCGAGCGATGTCGGCCCCCGGATAATATCGTTGCGCAAG 3916

RESULT 9

US-10-415-562A-7/c
; Sequence 7, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CL1709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 9652
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-415-562A-7

Alignment Scores:

Pred. No.: 1,28e-214 Length: 9652
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-415-562A-7 (1-9652)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20

Db 5052 ATGACGAGCGTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCGCTCGTGTGC 4993
Qy 21 SerAspLysArgGlyValLeuArgHisGluLeuArgProLysLeuGlnGlnThrThrSer 40
Db 4992 TCCGATTAAGCGCGGCAATCCGACGAACTGCGACCAAACTTCAACAAATCACACGTC 4933
Qy 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 4932 GAAACATTTAAACGCTGTGGCGCGCGGATTTCTGGCGTGAACGGGTGACCATTTGTCAAC 4873
Qy 61 GlyProLysGlySerGlyPheGlyValLeuArgSerCysGlyLysGlyValThrIleCysPro 80
Db 4872 GGTCCGAAGGTTCTGGATTCTGGAGGCTTTCTTCTCGGAAAGGCTGATCTGCC 4813
Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 4812 TGCTGTGGGGAAGTCTGGTGCATCTGTCAGACGAAATTTCTCAAGTTGTGCTCAT 4753
Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 4752 CAACTCGGAGCTGATCTGTTGCGATGTTGACGATGACCATGCGCCATACAGCTGGTCAG 4693
Qy 121 ArgIleuHisAspLeuThrGlyLeuSerAlaAlaTTPLeuLysAlaAlaThrAsnGlyArg 140
Db 4692 CGGCTCCACGACCTATGGAATGGAATTTCTCGGAGCTGGAAGCTGCGACCAACGGTCTGT 4633
Qy 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 4632 CGTTGGGCTACGGAAGTGAATGATGACGCTCGGACGATACGTCGGGCTGTGAAATC 4573
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 4572 ACTCACGGAAAAACCGCTGGACGCTCCACGTTCCAGCGCTACTCATGTTCTAGTGTGAC 4513
Qy 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db 4512 GTGAGTGAAGAAATCTCGAATCTTCTCGATGCGATGTTGATCGGTGGATCTGCAAA 4453
Qy 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyCysLeuAspValArgLys 220
Db 4452 CTCGTATCTCTGGATTGCTGGCCACTACGTAATTCGGTGGTCTCGATGTACGAAAG 4393
Qy 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrIleuThrLysIleAlaSerGlyVal 240
Db 4392 ATCGGCGGTGAAGCTGATCAAGTCTCTCGTGGTATCTGACGAAATTTGATCTGGCGIT 4333
Qy 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 4332 GGTATGGAGTTGTTAGTGGCGACGGAAGAAAGTGTGACATGCGCAACCGTGCACCTGG 4273
Qy 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 4272 GAAATCGCTGTTGATGAGTGGCGGGATCCCAAGCGTTGAACTGTGGCGAGAATTT 4213
Qy 281 GluPheGlySerMetGlyArgAlaAlaIleAlaTrpSerArgGlyLeuArgAlaProAla 300
Db 4212 GAGTTTGGTTGATGGAGCTCGGCAATCGGTGGTCCGTTGGATGGCTGGCCGAGCT 4153
Qy 301 GlyLeuGlyValAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 4152 GGTCTTGGGCGAGATTCACAGATGCTGATGATGATGATGATGATGATGATGATGATGATG 4093
Qy 321 MetValAlaIlelleProAlaArgSerTrpMetMetIleArgThrCysAlaProTyrVal 340
Db 4092 ATGGTTGCGATCATTCGCGCGCGATCGTGGATGATGATGATGATGATGATGATGATGATG 4033
Qy 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 4032 TTCGGCGAGATCTCTCGACCTCGTGAAGCTGGCGGACTTGGGAAATCTTCGGATCAC 3973
Qy 361 LeuHisTyrArgLeuProAlaAlaAspValArgProIlelleSerValArgLys 379

Db 3972 TTGCATTATCGATTGCCCCGAGCGGATGTGCGGCCCCCGATATATCGGTTGCGCAAG 3916
RESULT 10
US-10-007-527A-6/c
; Sequence 6, Application US/10007527A
; Publication No. US20030044807A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,527A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-527A-6
Alignment Scores:
Pred. No.: 1,55e-214 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1.4 Gaps: 0
US-10-007-527A-2 (1-379) x US-10-007-527A-6 (1-11241)
Qy 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
Db 6641 ATGACGAGCGTAAGTGTGAACACCTTTCCGGCAAGACCGGCTCCGCTCGTGTGC 4582
Qy 21 SerAspLysArgGlyValLeuArgHisGluLeuArgProLysLeuGlnGlnThrThrSer 40
Db 6581 TCCGATTAAGCGCGGCAATCCGACGAACTTCCGACCAAACTTCAACAAATCACACGTC 6522
Qy 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 6521 GAAACATTTAAACGCTGTGCGCGCGGATTTCTGCGCTGCAACGGTGTGACCATTTGTCAAC 6462
Qy 61 GlyProLysGlySerGlyPheGlyLeuArgSerCysGlyLysGlyValThrIleCysPro 80
Db 6461 GGTCCGAAAGGTTCTGGATTCTGGAGGCTTCTGTTCTGCGAAAGGCTGGATCTGCC 6402
Qy 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 6401 TGCTGTGGGGAAGTCTGGTGCATCTGTCAGACGAAATTTCTCAAGTTGTGCTCAT 6342
Qy 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 6341 CAACTCGGAGCTGATCTGTTGCGATGTTGACGATGACCATGCGCCATACAGCTGTGTCAG 6282
Qy 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 6281 CGGCTCCACGACCTATGAGCTGGACTTTCCGACGCTTGGAAAGCTGGAAAGCTGCGCAACGGTCTGT 6222
Qy 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 6221 CGTTGGCGTACGGAACGTAAGATGTACGGCTGCGACGATACGTCGGCGCTGTGTAATC 6162
Qy 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 6161 ACTCACGGAAAAACCGCTGGACGCTCCAGCTTCCAGCGCTACTCATGTTCTGAGTGTGAC 6102
Qy 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200

Db 6101 GTGAGTGAAGACATCTCTCGAATCTTCTCGATGCGATGTTTCGATCGGTGGACTTCCAAA 6042
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 6041 CTGCTATCTCTGGATTGCTGCGCCACTACGTAAATTCGGGTGTTCTCGATGTACGAAAG 5982
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 5981 ATCGCGGTGAAGCTGATCAAGTCTCTCGTTCGTATCTGACGAAATTCATCTGGGTT 5922
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 5921 GGTATGGAGGTGGTATGTCGACGGAAGAGTGGTCGACATGCAACCGTGCACTCTG 5862
QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 5861 GAAATCGCTGTGATCGATGGCGGGGATCCAAAGCGTTGGAATCTGGCGAGAATTT 5802
QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 5801 GAGTTTGGTTCGATGGAGCTCGGGCAATCGGTGGTCCGTGATGGTGGCCGAGCT 5742
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGluGluSerAlaProVal 320
Db 5741 GGTCTTGGGGCAGAACTAACAGATGCTCAGATGCTTGACGAGCAAGAACTCTGCCCGGTC 5682
QY 321 MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTrpVal 340
Db 5681 ATGGTTGGATCATTTCCGGCGGATCGTGGATGATTCGGACTTGTGCGCTTACGTC 5622
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 5621 TTCGGCGAGATCCTCGGACTCGTCGAAGCTGCGGACTTGGGAAATCTTCGTGATCAC 5562
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleSerValArgLys 379
Db 5561 TTGCATTATCGATTCCCGCAGCGATGTGCGGCCCGCGATAATATCGTTTCGCAAG 5505

RESULT 11

US-10-007-452-6/c
; Sequence 6, Application US/10007452
; Publication No. US20030093701A1
; GENERAL INFORMATION:
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Bramucci, Michael G.
; APPLICANT: Cheng, Qiong
; APPLICANT: Kostichka, Kristy N.
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: CLI709 US NA
; CURRENT APPLICATION NUMBER: US/10/007,452
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-007-452-6

Alignment Scores:

Pred. No.: 1,55e-214 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-007-527A-2 (1-379) x US-10-007-452-6 (1-11241)

QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProProValLeuValSer 20
|||||

Db 6641 ATGACCAGCGTAAGTCTGAACACCTTTCGCGCAAAAGACCGGCTCTCCGCTCTCTGTCG 6582
QY 21 SerAspLysArgGlyIleArgHisGluLeuArgProLysLeuGlnIleThrThrSer 40
Db 6581 TCCGATAAGCGCGCATCCGGCACGAAGCTGCGACCAAACTTCAACAAATTCACACGCTCA 6522
QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrIleValAsn 60
Db 6521 GAAACATTTAAACCCCTGTGCGCGCGGATTTCTGGCGTGAACGGTGTGACCATTTGTCAAC 6462
QY 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyTrpIleCysPro 80
Db 6461 GGTCCGAAAGGTCTCGATTGCGAGGCTTCTGCTCGGAAAGGCGCTGGATCTGCCCC 6402
QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnValValAlaHis 100
Db 6401 TGCTGTGCGGAAAAGTCTGGTGCACATCTGTCAGACGAAATTTCTCAAGTTGTGTCTCAT 6342
QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGln 120
Db 6341 CAACTCGGAGCTGGATCTCTGTCGATGCTGACCATGACCATGCGCCATACAGCTGCTCAG 6282
QY 121 ArgLeuHisAspLeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 6281 CGGCTCCAGACCTATGGACTGGACTTTCGGCAGCTCGGAAAGCTCGGACCAACGCTCGT 6222
QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIle 160
Db 6221 CGTGGCGTACGGAACGTGAAATGTACGGCTCGCAGCGGATACGTGGCGCTGTGAAATC 6162
QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 6161 ACTCAGGAAAGAAACCGCTGCGCACGCTCCACGTTCCACGCGTACTCATGTTCAAGTGGTAC 6102
QY 181 ValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThrSerLys 200
Db 6101 GTGAGTGAGAACATCCTCGAATCTTCGATGCGATGCTTCGATCGGTGGACTTCCAAA 6042
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220
Db 6041 CTCGTATCTCTGGGATTGCTCGGCCACTACGTAAATTCGGTGGTCTCGATGTACGAAAG 5982
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 5981 ATCGCGGTGAAGCTGATCAAGTTCCTCGCTGCTATCTGACGAAATTCATCTCGGTT 5922
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 5921 GGTATGGAGGTGGTATGTCGCGCGGATCCACACGCTTGGAACTGTGGCGAGAATTT 5802
QY 261 GluIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 5801 GAAATCGCTGTGATGTCAGTGGCGGGGATCCACACGCTTGGAACTGTGGCGAGAATTT 5742
QY 281 GluPheGlySerMetGlyArgArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 5741 GGTCTTGGGCGAAGACTTAAACAGATGCTGATGCTGAGCAGGAAAGAAATCTGCGCCGCTC 5682
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGluGluSerAlaProVal 320
Db 5681 GGTCTTGGGCGAAGACTTAAACAGATGCTGATGCTGAGCAGGAAAGAAATCTGCGCCGCTC 5622
QY 321 MetValAlaIleProAlaArgSerTrpMetIleArgThrCysAlaProTrpVal 340
Db 5621 GGTCTTGGGCGAAGACTTAAACAGATGCTGATGCTGAGCAGGAAAGAAATCTGCGCCGCTC 5562
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 5561 TTCGGCGAGATCCTCGGACTCTGCGAGCTGCGGAGCTTGGGAAATCTTCGTGATCAC 5505
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProProIleSerValArgLys 379
Db 5561 TTGCATTATCGATTCCCGCAGCGATGTGCGGCCCGCGATAATATCGTTTCGCAAG 5505

RESULT 12
US-10-415-562A-6/c
; Sequence 6, Application US/10415562A
; Publication No. US20040115661A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Rhodococcus Cloning and Expression Vectors
; FILE REFERENCE: C11709 US PCT
; CURRENT APPLICATION NUMBER: US/10/415,562A
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/254,868
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 11241
; TYPE: DNA
; ORGANISM: Plasmid pRHBR17
US-10-415-562A-6
Alignment Scores:
Pred. No.: 1,55e-214 Length: 11241
Score: 2005.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-10-007-527A-2 (1-379) x US-10-415-562A-6 (1-11241)
QY 1 MetThrSerValSerAlaGluHisLeuSerGlyLysAspArgProValLeuValSer 20
Db 6641 ATGACGAGCGTAAAGTCTGAACACCTTTCCGGCAAAAGACCGGCTCCCGTCTCGTGC 6582
QY 21 SerAspLysArgGlyThrLeuArgHisGluLeuArgProLysLeuGlnLeuThrThrSer 40
Db 6581 TCCGATAACGCGGCGATCCGGCAGCAACTGCGACCCCAAACTTCAACAAATCCACGTC 6522
QY 41 GluThrPheAsnAlaCysGlyArgProLysSerGlyValAsnGlyValThrLeuValAsn 60
Db 6521 GAAACATTAAACGCTGTGCGCGCGATTTCTGGGTGACGATGACCGTGTGACCATTTGTC 6462
QY GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGlyThrLysCysPro 80
Db 6461 GGTCCGAAAGGTTCTGGATTCCGAGGCGCTTCTGTTCTCGGAAAGGCGTGGATCTCC 6402
QY 81 CysCysAlaGlyLysValGlyAlaHisArgAlaAspGluLysSerGlnValValAlaHis 100
Db 6401 TGCTGTGCGGAAAGTCGGTGCACATCGTGCAGCAAAATTTCTCAAGTTGTGTCTCAT 6342
QY 101 GlnLeuGlyThrGlySerValAlaMetValThrMetThrArgHisThrAlaGlyGln 120
Db 6341 CAACTCGGACGTGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6282
QY 121 ArgLeuHisAspLeuThrThrGlyLeuSerAlaAlaThrLysAlaAlaThrAsnGlyArg 140
Db 6281 CGGCTCCACGACCATATGAGTGGATGAGTTCGCGAGCCTGGAAGCTGCGACCAACGTCGT 6222
QY 141 ArgThrArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluLeu 160
Db 6221 CGTGGCGTACGGAAGTGAATGTACCGGTGCGGATGATGATGATGATGATGATGATGAT 6162
QY 161 ThrHisGlyLysAsnGlyTyrPheHisValHisAlaLeuLeuMetPheSerGlyAsp 180
Db 6161 ACTCAGGAAACGCGTGGCAGCTCCACGTTTCCGCGCTACTCATGTTTCAGTGTGTGAC 6102
QY 181 ValSerGluAsnLeuLeuGluSerPheSerAspAlaMetPheAspArgThrThrSerLys 200
Db 6101 GTGAGTGAGAACATCTCTCGAATCTCTCGAATCTCTCGAATCTCTCGAATCTCTCGAAT 6042
QY 201 LeuValSerLeuGlyPheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLys 220

6041 CTCGTATCTCTGGGATTTGTCGCCACTAGTAATTCGGGTGTCGTGATGTAAGAAAG 5982
QY 221 IleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal 240
Db 5981 ATCGCGGTGAAGCTGATCAAGTTCTCGTCCGTAATCTGACGAAATTCATCTGGCGTT 5922
QY 241 GlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrp 260
Db 5921 GGTATGAGGTGTGTAGTGGCGAGGAAAGTGTGCATGTCACACCGTGACCCCTGG 5862
QY 261 GluIleAlaValAspAlaValGlyAspProGlnAlaLeuGluLeuTrpArgGluPhe 280
Db 5861 GAAATCGCTGTGATGTCAGTGGCGGATCCACAGCGTTGGAACTGTGGCGAAATTT 5802
QY 281 GluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaArgAla 300
Db 5801 GAGTTTGGTTCGATGGGACGTCGGGCAATCCCGTGTCCCTGGATTGCGTCCGAGCT 5742
QY 301 GlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaProVal 320
Db 5741 GGTCTTGGGCGAGAACTTAACAGATGCTCAGATCGTTGAGCAGGAAGAATCTGCCCGGTC 5682
QY 321 MetValAlaIleIleProAlaArgSerTrpMetMetIleArgThrCysAlaProTrpVal 340
Db 5681 ATGTTTGGATCATTTCCGGCGGATCGTGGATGATGATGATGATGATGATGATGATGAT 5622
QY 341 PheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTrpGluAsnLeuArgAspHis 360
Db 5621 TTCGGCGAGATCTCGGACTCTCGAAGCTGGCGGCACTTGGGAAATCTTCGTGATCAC 5562
QY 361 LeuHisTyrArgLeuProAlaAlaAspValArgProLysLeuSerValArgLys 379
Db 5561 TTGCATTATCGATTGCCCGCAGCGATGTGCGGCCCGGATAATATCGTTGCGCAAG 5505
RESULT 13
US-09-754-112A-2/c
; Sequence 2, Application US/09754112A
; Publication No. US20020090687A1
; GENERAL INFORMATION:
; APPLICANT: TRIPATHI, Gyanendra, et al.
; TITLE OF INVENTION: Processes for the Production of Polyhydroxyoctanoate by Streptomyces
; FILE REFERENCE: 041144.006
; CURRENT APPLICATION NUMBER: US/09/754,112A
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Streptomyces lividans
US-09-754-112A-2
Alignment Scores:
Pred. No.: 5.9e-27 Length: 7560
Score: 337.00 Matches: 107
Percent Similarity: 42.46% Conservative: 45
Best Local Similarity: 29.89% Mismatches: 126
Query Match: 16.81% Indels: 80
DB: 9 Gaps: 16
US-10-007-527A-2 (1-379) x US-09-754-112A-2 (1-7560)
QY 23 LysArgGlyThrLeuArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSerGluThr 42
Db 1605 CGAAAGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1546
QY 43 PheAsnAlaCysGlyArg-----ProfileSerGlyVal----- 53
Db 1545 TGCAAGGCGTGGCGCGCGCGCTGATGACCTCTGCTCCGCGTGTGATGTCGCGCCAGACG 1486
QY 54 ---AsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSer 72

```

Db 1485 GCGCGCGGAACGTCCTGTCCTC-----GGCTGATGCGG 1450
QY 73 CysGlyLysGlyTyrPileCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
Db 1449 TGGCGCGGATCTGGCTCTCCCGGCTGCGCGCGCCACGATCCGCGCAACAAGCGGCGGAG 1390
QY 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAla---MetValThr 111
Db 1389 GAGATCACCGCCCGCGTGTGAGTGGATCAACGCGGGGGGACCGCTACCTGGTCACC 1330
QY 112 MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrPheGlyLeuSerAla 131
Db 1329 TTCACGCGCGCGCATGGGACACGCGCGCTCGCGGACCTCATGACGCGCTCCAGGGC 1270
QY 132 AlaTrpLysAlaAlaThrAsnGlyArg-----141
Db 1269 ACCCGAAGACGCGGACAGCCCGCGCGCGCGCTTACGACGACTGATCACGGG 1210
QY 142 -----TrpArgThrGlu-----145
Db 1209 GGCACGTGGCGCGGACCGCGCGCCAAAGACGCGGCGCGCGCGCGCGCGCGCGCATC 1150
QY 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLys---164
Db 1149 CGAGACCGGATCGGGTACGTCGCGATGATCCGCGGACCGAAGTCAACCGTGGCGCAGATC 1090
QY 165 AsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal-----181
Db 1089 AACGCTGGCACCGCGACATCACCGGATCGTGTGCGGCGCGCGCGCGCGCGCGCGGAG 1030
QY 182 -----SerGluAsnIleLeuGluSerPhe-----141
Db 1029 CGGTCCCGAAGACGATGCTGCGCCACCTTCAGCGCGACCGCGCGCGCGCGCGCGGAG 970
QY 198 ThrSerLysLeuValSerLeuGlyPheAlaProLeuArg-----211
Db 969 CAGCGGCACTGGCGGTCCTG-----TGGACCGCGCGCGCGCGCGCGCGCGCGCGCTTC 913
QY 212 -----AsnSerGlyLysLeuAspValArgLysIleGlyGlyGluAlaAsp---Gln 227
Db 912 ACGCCCGACACCGCGCGCGCGCTCGACTTCAAGCGCGTGGAGACCGCGCGCGCGCGCGG 853
QY 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle-----262
Db 792 CGCGCGGACCTCAAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 733
QY 263 AlaValAspAlaValGlyGly-----AspProGlnAla-----273
Db 732 ATCGGGGACCTGACCGCGCGCATACCGAGGACGACCGCGCGCGCGCGCGCGCGCGCGCG 673
QY 274 -----LeuGluLeuTrpArgGluPheGlySerMetGlyArgArgAlaIleAla 291
Db 672 TGGAACTCTCGCGTGGCAGAGTACGAGCGGGCAACCGCGGGGACCGCGCGCGCGCGCGG 613
QY 292 TrpSerArgGlyLeuArgAlaAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCGCGCTACCTGCGCGAGATGCTGGGCTCGACGCGCGCGCGCGCGCGCGCGG 559

```

RESULT 14

```

US-09-754-112A-1/c
; Sequence 1, Application US/09754112A
; Publication No. US20020090687A1
; GENERAL INFORMATION:
; APPLICANT: TRIPATHI, Gyanendra, et al.
; TITLE OF INVENTION: Process for the Production of Polyhydroxyoctanoate by Streptomyces
; TITLE OF INVENTION: lividans
; FILE REFERENCE: 041144.006
; CURRENT APPLICATION NUMBER: US/09/754,112A
; CURRENT FILING DATE: 2001-01-05

```

```

; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 12738
; TYPE: DNA
; ORGANISM: Streptomyces lividans
US-09-754-112A-1

Alignment Scores:
Pred. No.: 1,15e-26      Length: 12738
Score: 337.00           Matches: 107
Percent Similarity: 42.46%      Conservative: 45
Best Local Similarity: 29.89%    Mismatches: 126
Query Match: 16.81%            Indels: 80
DB: 9                      Gaps: 16

US-10-007-527A-2 (1-379) x US-09-754-112A-1 (1-12738)

```

```

QY 23 LysArgGlyIleArgHisGluLeuArgProLysLeuGlnGlnIleThrThrSerGluThr 42
Db 1605 CGGAAGCGCGCGCTGACAGCGCGCGCGCTGTGGCGGATCAGCGGGGACGCGCGG 1546
QY 43 PheAsnAlaCysGlyArg-----ProIleSerGlyVal-----53
Db 1545 TCACAGGCGTGGCGCGCGCGCGCTGATGACCTGCTCGCGGCTGATCGTCCCGACAGC 1486
QY 54 ---AsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeuArgSer 72
Db 1485 GCGCGCGGAACGTCGTCGTCTG-----GGCTGATGCGG 1450
QY 73 CysGlyLysGlyTyrPileCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAsp 92
Db 1449 TGGCGCGGATCTGGCTCTGCGCGCTGCGCGCGCACGATCCGCGCACAAAGCGGCGGAG 1390
QY 93 GluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAla---MetValThr 111
Db 1389 GAGATCACCGCGCGCTGTCAGTGGATCAAGCGCGGCGCGCGCGCGCGCGCTACCTGTCACC 1330
QY 112 MetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSerAla 131
Db 1329 TTCACGCGCGCGCATGGGACACGCGCGCTCGCGGACCTCATGAGACCGCGCTCCAGGGC 1270
QY 132 AlaTrpLysAlaAlaThrAsnGlyArg-----141
Db 1269 ACCCGAAGACGCGGACAGCCCGCGCGCGCGCGCGCTTACGACGACTGATCACGGCG 1210
QY 142 -----TrpArgThrGlu-----145
Db 1209 GGCACGTGGCGCGGACCGCGCGCCAAAGACGCGGCGCGCGCGCGCGCGCGCGCATC 1150
QY 146 ArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLys---164
Db 1149 CGAGACCGGATCGGGTACGTCGCGATGATCCGCGCGACCGAAGTCAACCGTGGCGGAGATC 1090
QY 165 AsnGlyTyrHisValHisValHisAlaLeuLeuMetPheSerGlyAspVal-----181
Db 1089 AACGCTGGCACCGCGACATCACCGGATCGTGTGCGCGCGCGCGCGCGCGCGCGGAG 1030
QY 182 -----SerGluAsnIleLeuGluSerPhe-----141
Db 1029 CGGTCCCGAAGACGATGCTGCGCCACCTTCAGCGCGACCGCGCGCGCGCGCGCGGAG 970
QY 198 ThrSerLysLeuValSerLeuGlyPheAlaProLeuArg-----211
Db 969 CAGCGGCACTGGCGGTCCTG-----TGGACCGCGCGCGCGCGCGCGCGCGCGCGCTTC 913
QY 212 -----AsnSerGlyLysLeuAspValArgLysIleGlyGlyGluAlaAsp---Gln 227
Db 912 ACGCCCGACACCGCGCGCGCGCTCGACTTCAAGCGCGTGGAGACCGCGCGCGCGCGCGG 853
QY 228 ValLeuAlaAlaTyrLeuThrIleAlaSerGly-----ValGlyMetGluValGly 245
Db 852 GACCTCGCGGATGATCGCCAAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 793

```

QY 246 SerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIle----- 262
Db 792 CGCGCCGACCTCAAGACGGCGGACCGGAGACGTCGCCCGTTTCGAACCTCCTCGGACGG 733
QY 263 AlaValAspAlaValGlyGly-----AspProGlnAla----- 273
Db 732 ATCGGGACCTTGACCGCGGATGACCCGAGACGACCGCGCGGGTCCGCTCGCTGGAG 673
QY 274 -----LeuGluLeuTrpArgGluPheGlySerMetGlyArgAlaIleAla 291
Db 672 TGGAACTCTCGCGTGCACAGATACGAGCGGGCAACCGCGGAGCGCGGCCATCGAA 613
QY 292 TrpSerArgGlyLeuArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 612 TGGACCCGCTACCTGCGGCAGATGCTCGGGCTCGACGCGCGGCGACACCGAGGCC 559

RESULT 15

US-09-835-381-5
; Sequence 5, Application US/09835381
; Patent No. US20020045223A1
; GENERAL INFORMATION:
; APPLICANT: SUGA, Mikiko
; APPLICANT: ASAKURA, Yoko
; APPLICANT: MORI, Yukio
; APPLICANT: ITO, Hisao
; APPLICANT: KURAHASHI, Osamu
; TITLE OF INVENTION: ARGININE REPRESSOR DEFICIENT STRAIN OF CORYNEFORM BACTERIUM AND METHOD FOR PRODUCING L-ARGININE
; FILE REFERENCE: 206018US0
; CURRENT APPLICATION NUMBER: US/09/835,381
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: JP2000-129167
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 4447
; TYPE: DNA
; ORGANISM: Brevibacterium lactofermentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1318)..(2598)
; OTHER INFORMATION:
US-09-835-381-5

Alignment Scores:
Pred. No.: 5, 19e-26 Length: 4447
Score: 326.00 Matches: 101
Percent Similarity: 47.75% Conservative: 58
Best Local Similarity: 30.33% Mismatches: 133
Query Match: 16.26% Indels: 42
DB: 9 Gaps: 14

US-10-007-527A-2 (1-379) x US-09-835-381-5 (1-4447)

QY 34 LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgProIleSerGlyVal 53
Db 1396 ATGTACAAGATCAACCAATAGTAGGCTCTGGCGGGGTGCCATAGGTG-GCGCAGGACGA 1454
QY 54 Asn-----GlyValThrIleVal-AsnGlyProLysGlySerGlyPheGlyLeuAr 71
Db 1455 AGCTGTGGGTGTCTCTGGTCTGCTAACGGT-----GCTTCGAGTTTTCAGGCTCGCA 1508
QY 71 gSerCysGlyLysGlyTrpIleCysProCysAlaGlyLysValGlyAlaHisArgAl 91
Db 1509 AAACCTCACTCTCGCTGGGGGTCACTCTGGTGAATTCGAAGTCAATGCGGAAACGCCG 1568
QY 91 AspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAlaMetValTh 111
Db 1569 CATTGAGCTGGCTATTCTACTAAGAAATCACTTGGCGCGGGTGGCGCGTTCATGATGTT 1628
QY 111 rMet---ThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeuSe 130

Db 1629 TGTGGGCACTGTTTCGACACCAACCGCTCACAGTCATTTTCGCGAGGTTGAAGCGGTATTAA 1688
QY 130 rAlaAlaTrpIlysalalalaThrAsnGlyArgGlyTrpArgThrGluArgGluMetTyxGl 150
Db 1689 GACTCGCTACTCTTCGATGTGTGAACCATCTCATGTGAAGAAGAAACCTGCACCGTACGG 1748
QY 150 yCysAspGlyTyxValArgAlaValGluIleThrHisGly---LysAsnGlyTrpHisVa 169
Db 1749 GGTGGACACACCTATAGTACATAGGTACACAGACTCTTGGGCGCAACGGTTGGCACTT 1808
QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPh 189
Db 1809 GCACCGCAACATGCTGTTTCTTGGATCGTCCACTGTCTGACATGAACACTCAAGCGCGTT 1868
QY 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 1869 TGAGGATTCCATGTTTTCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1928
QY 209 oLeuArgAsnSerGly---GlyLeuAsp---ValArgLysIleGlyGlyGluAlaAspGl 227
Db 1929 ACTCGGTGACGACGGGGTCAAACTTGATCAGGTGTCTACCTGGGGTGGAGACGCTGCGAA 1988
QY 227 nValLeuAlaAlaTyxLeuThrLysIleAlaSerGlyValGlyMetGluValGlySerGl 247
Db 1989 A---ATGGCAACCTACCTCGCTAAG-----GGCATGCTCTCAGGAACCTGCTGGCTC 2036
QY 247 yAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaVa 267
Db 2037 CGCTACTAAACCGCGTCTTAAGGGTCTGATACGCGGCTTTCAGATG---TTGGATATGTT 2093
QY 267 lGlyGlyAspProGlnAlaLeuGlu-----LeuTrpArgGl 279
Db 2094 GGCCGATCAAGCGGCGCGGAGATATGGAGCTGTTTGGTGGCTCGGTCGCGTGA 2153
QY 279 uPheGluPheGlySerMetGlyArgAlaIleAlaTrpSerArgGlyLeuArgAlaAr 299
Db 2154 GTATGAGGTGGTCTTCTAAACCTCGCTCG---TCCTGGTCACTGGG-----GCTAA 2204
QY 299 gAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIle-----ValGluGlnGlu--- 315
Db 2205 GCGTGTCTTGGGCACTTGATTACATAGCTGATGATGCTGTAATGGAAGAACT 2264
QY 316 -----GluSerAlaProValMetValAl 323
Db 2265 GTACAAGCTCGCGGTCTGGAAGCACCGGTCGAATCAACCGCGGTTGCTGTGTC 2324
QY 323 alleleProAlaArgSerTrpMetMetIleArgThr 335
Db 2325 TTTGTGAAGCCCGCATGATTGGAACACTGATTCAGTCT 2361

Search completed: October 23, 2004, 08:18:11
Job time : 659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 03:42:56 ; Search time 24 Seconds
(without alignments)
1519.423 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGKDRPPVLVS.....HLHYRLPAADVRPPIISVRK 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: Piri: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321.5	16.0	256	2 S34850	hypothetical prote
2	316	15.8	456	2 A31844	rep protein - Stre
3	230.5	11.5	437	2 S04020	hypothetical prote
4	169	8.4	240	2 I40335	hypothetical prote
5	127	6.3	314	2 B35390	replication protei
6	123.5	6.2	315	2 S25721	replication protei
7	122.5	6.1	315	2 T44800	replication protei
8	118.5	5.9	315	2 JN0856	rep protein - Str
9	115	5.7	336	2 A40651	replication protei
10	114	5.7	314	2 T46831	plasmid replicatio
11	111.5	5.6	333	2 A49783	replication initia
12	105	5.2	381	2 A83561	probable type II s
13	104.5	5.2	340	2 I40561	replication initia
14	104	5.2	420	2 AG2059	heme biosynthesis
15	102.5	5.1	420	2 AC3284	D-amino-acid dehyd
16	101.5	5.1	339	2 S01098	rep protein - Bac
17	100.5	5.0	1443	2 G75393	hypothetical prote
18	99.5	5.0	614	2 D87410	old protein (impo
19	98.5	4.9	532	2 H83493	hypothetical prote
20	98.5	4.9	3535	2 B83641	probable hemagglut
21	98	4.9	442	2 D87594	genetic exchange p
22	95.5	4.8	606	2 D83484	probable acyl-CoA
23	94.5	4.7	1198	2 T28678	polyketide synthas
24	94.5	4.7	1763	2 T17465	rifamycin polyketi
25	94	4.7	317	2 A43621	probable replicati
26	94	4.7	382	2 AE2861	l-lactate dehydrog
27	94	4.7	382	2 D97638	l-lactate dehydrog
28	94	4.7	544	1 BVC06L	chaperonin groEL
29	94	4.7	1245	2 H87340	hypothetical prote

30 93 4.6 317 2 S25294 replication protei
31 93 4.6 748 2 T37097 probable secreted
32 93 4.6 796 2 S57844 lethal(3) malignant
33 92.5 4.6 354 2 B75355 hypothetical prote
34 92.5 4.6 600 2 E72027 phosphoenolpyruvat
35 92.5 4.6 600 2 A85597 phosphoenolpyruvat
36 92.5 4.6 1026 2 G87346 hypothetical prote
37 92 4.6 799 2 T48889 serine/threonine p
38 91.5 4.6 302 2 H96792 unknown protein F1
39 91.5 4.6 332 2 S51694 replication protei
40 91.5 4.6 350 2 B82281 ferric vibriobacti
41 91 4.5 296 1 OPNE7 peroxidase (EC 1.1
42 91 4.5 356 2 B83200 probable phosphotr
43 90.5 4.5 320 2 D96750 unknown protein F2
44 90.5 4.5 2399 2 AH3009 non-ribosomal pept
45 90.5 4.5 2399 2 F98274 hypothetical prote

ALIGNMENTS

RESULT 1

S34850
hypothetical protein - Streptomyces lavendulae plasmid pSLG33
C:Species: Streptomyces lavendulae
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 07-May-1999
C:Accession: S34850
R:Felberg, J.; Petricek, M.; Tichy, P.
Nucleic Acids Res. 21, 3582, 1993
A:Title: Nucleotide sequence of the mini-plasmid pSLG33 from Streptomyces lavendulae-g
A:Reference number: S34850; MUID:93348001; PMID:8346038
A:Accession: S34850
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-256 <PEL>
A:Cross-references: EMBL:X69872
C:Genetics:
A:Genome: plasmid pSLG33

Query Match 16.0%; Score 321.5; DB 2; Length 256;
Best Local Similarity 33.2%; Pred. No. 8e-19;
Matches 85; Conservative 45; Mismatches 105; Indels 21; Gaps 11;

QY 109 MVTMTHTAQRHLHDLWTGLSAWKAATNGRWR-----TERMYGCDGYVRAVEI 160
Db 1 NVTLTARHKKHHELEPLFDVANGWRKLLSGAWGDPKRGVLGERDLGVGNIRSLV 60
QY 161 THG-KNGWHVHVHALLMFSGDVSNILESFSDAMFDR-WTSKLVSGLGFAAPLRNSGGLDV 218
Db 61 TYGRNGWPHLHVLLNNEETTEL--AYAMHWDKTRAWLKKAGFEPSKEH--GITW 116
QY 219 RKIGGEADQVLAAYLTIKIASG-VGMEVSGDGKSGRHNAPWE-IAVDVGGDPQALEL 276
Db 117 SKV--TTPPEAGEVIAKAQEGKVGNEIARGDKRAGLTGLTAPFEMLEYFRQTGMVVP 174
QY 277 WREFPGSGRRATJAWSGRLAR-AGLGAEITDAQIVEQESAPVWVAIIPARSMWIRT 335
Db 175 WQEEKGTFRRAITWRSRLRAELGLDEELTDELAEEIGGETW-ALLPAESLRAIRR 233
QY 336 CAPVFGIILGLVEAG 351
Db 234 -VFGQSRILDTAENG 248

RESULT 2

A31844
rep protein - Streptomyces lividans plasmid pIJ101
C:Species: Streptomyces lividans
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C:Accession: A31844
R:Kendall, K.J.; Cohen, S.N.
J. Bacteriol. 170, 4634-4651, 1988
A:Title: Complete nucleotide sequence of the Streptomyces lividans plasmid pIJ101 and c

A:Reference number: A91888; MUID:89008081; PMID:3170481
A:Accession: A31844
A:Molecule type: DNA
A:Residues: 1-456 <KEN>
A:Cross-references: UNIPROT:P22406
C:Genetics:
A:Genome: plasmid

Query Match 15.8%; Score 316; DB 2; Length 456;
Best Local Similarity 30.5%; Pred. No. 4.4e-18;
Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

QY 49 PISGV-----NGVTIVNGPKSGFGGLRSCGKMICPCCKAGKVAHRADEISQVVAQL 102
DB 3 PASGVIVAQAAGTSVVL-----GLMRCGRILWCPVCAATIRHKRAEIEITAAVVEVI 54

QY 103 GTGSAV-MVTMTMRTAQRLHDLWTGLSAWKAATNGRR-----WRTE----- 145
DB 55 KRGGTAYILVTFARHGHTDRLADLMDALQGTTRTPDPRPGAYQRLITGGTWAGRAKD 114

QY 146 -----REMYGCDGYRAVEITHGK-NGWHVHVHALLMFSGDV-----SENILESF 189
DB 115 GHRAADREGIRDIGVGMIRATEVTVGQINGWHPHIHAIVLVGKRTGERSAKQIVATF 174

QY 190 --SDAMFDRWTSKLVSLGFAAPLR-----NSGLDVRKIGIGHEAD-QVLAAYLTAKIA 237
DB 175 EPTGAALDEWQHWRSV-WTAALRKVNPAFTPDRHGVDFKRLETERDANDLAELYIAKQ 233

QY 238 SG--VCMFVSGDGKSGRGNAPWEI---AVDVG---DPOA-----LELWREPE 281
DB 234 DGKAPALEARADLTKATGNAVPELLOGLIGLGTGMDTDAAGVSGLENNLSRWHEYE 293

QY 282 FGSMDGRRALAWSGRLRARAGLGAELTDA 309
DB 294 RATRGRRALEWRYRLQMLGLDGDTEA 321

RESULT 3
S04020
hypoetical protein - Streptomyces sp. plasmid pSB24.2
C:Species: Streptomyces sp.
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: S04020
R:Polotin, A.P.; Sorokin, A.V.; Alekandrov, N.N.; Danilenko, V.N.; Kozlov, Y.I.
Dokl. Biochem. 283, 260-263, 1985
A:Title: Nucleotide sequence of DNA of the actinomycete plasmid pSB24.2.
A:Reference number: S04020
A:Accession: S04020
A:Molecule type: DNA
A:Residues: 1-437 <BOL>
A:Cross-references: UNIPROT:Q52205; EMBL:X03756
A:Note: the authors translated the codon TGG for residue 59 as Arg, CGG for residue 187
A:Note: the source is designated as Streptomyces cyanogenus
C:Genetics:
A:Genome: plasmid

Query Match 11.5%; Score 230.5; DB 2; Length 437;
Best Local Similarity 24.6%; Pred. No. 4.1e-11;
Matches 102; Conservative 43; Mismatches 140; Indels 129; Gaps 19;

QY 55 GVTIVNGPKSGFG-GLRSCGKMICPCCKAGKVAHRADEISQVVAQLGTGSA-MVTM 112
DB 6 GVTFARTAAAGAAVALGKRIWLCVPVCSQDPAPFSEITEAVVSLQGGWAYILVF 65

QY 113 TMRTAQRLHDLWTGL-----SAWKAATNGRRMTEREM 149
DB 66 TARTAAADRLDMDALQGTTRADTETGKRPAYQRLITGAAM--AGDKRKNQEGIR 123

QY 150 GCDGY---VRAVEITHGKN-NGWHVHVHALLMFSGDV-----SENILESFS---DAMP--- 194
DB 124 GRIGYIGMIRATEVTVGAGWHPHIHAIVLVGKRTGERSAKQIVATF 183

195 DRWTSKLVSLGFAAPLRNSGGL-----DVRKIGIGHEADQVLAAYLTAKIASGVG----- 241
DB 184 DRWERS-----GPATLARSTPGFRPTTGARSPGATAGARATASTSSSVRSGRQRPGRVHR 239

QY 242 -----MEVSGDGKSGRHN-----RAPWETAIVAVGDDPQ-----ALELW 277
DB 240 QDAGRQEPGAGTRRPGQPPGQHDVLTFSRIG-DLMGVPPEEAAGHSGSLAWGLDRW 298

QY 278 REPEFGSMG-----RRATAWSRGL-----RARAGL 302
DB 299 AEYETAIVSGAGHRVDLPAPAGPDRRRHRGRRHGRPVPDRRRFRDGVQWIDRAWNGL 358

QY 303 CAELTDAQIVQEHSAPVWVAII-----PARSMMMIETCAPY-----VFEILG 346
DB 359 VGRSLDLAVVEAVEGREISMDALGELVQSAQSRAFLRLVLTPOEVTLEYDELLG 412

RESULT 4
I40335
hypoetical protein 1 - Corynebacterium glutamicum
C:Species: Corynebacterium glutamicum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I40335
R:Yamaguchi, R.; Terabe, M.; Miwa, K.; Tsuchiya, M.; Takagi, H.; Morinaga, Y.; Nakamori, A.; Title: Determination of the complete nucleotide sequence of Brevibacterium lactoferme
Agric. Biol. Chem. 50, 2771-2778, 1986
A:Title: Determination of the complete nucleotide sequence of Brevibacterium lactoferme
A:Reference number: I40334
A:Accession: I40335
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-240 <RES>
A:Cross-references: UNIPROT:Q45282; GB:D00038; NID:G216381; PIDN:BAA00026.1; PID:G216381

Query Match 8.4%; Score 169; DB 2; Length 240;
Best Local Similarity 32.0%; Pred. No. 2.2e-06;
Matches 54; Conservative 28; Mismatches 53; Indels 34; Gaps 10;

QY 193 MFDRTSKLVSLGFAAPLRNSG-GLD-VRKIGIGHEADQVLAAYLTAKIASGVMEVSGDGK 250
DB 1 MFSRWSAGVTVKAGMDAPLREHVKLDQVSTWGGDAK-MATYLAK---GMSQELTGSATK 56

QY 251 SCRHNRAPEIATAVAVGDDPQALE-----LWREFEFGSGRRALAWSGRLRARAGL 302
DB 57 TASKGSYTPFQM-LDMADQSDAGEDMDAVLVARWEYEVGSKNLS-SWSRG--AKRAL 112

QY 303 CAELTDAQI---VEQE-----ESAPVMVAIIPARSGMMIRT 335
DB 113 GDIYDADVRREMEELYKLAGLEAPERVESTVAVALKPPDDKLIQS 161

RESULT 5
B35390
replication protein - Lactobacillus hilgardii
C:Species: Lactobacillus hilgardii
C:Date: 31-Aug-1990 #sequence_revision 31-Aug-1990 #text_change 09-Jul-2004
C:Accession: B35390
R:Jooss, K.; Soetaert, P.; Michiels, F.; Joos, H.; Mahillon, J.
J. Bacteriol. 172, 3089-3099, 1990
A:Title: Lactobacillus hilgardii plasmid pLAB1000 consists of two functional cassettes
A:Reference number: A35390; MUID:90264300; PMID:2188951
A:Accession: B35390
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-314 <JOS>
A:Cross-references: UNIPROT:P35857; GB:M55222; NID:G149352; PIDN:AAA98163.1; PID:G149352

Query Match 6.3%; Score 127; DB 2; Length 314;
Best Local Similarity 22.0%; Pred. No. 0.0081;
Matches 55; Conservative 50; Mismatches 107; Indels 38; Gaps 10;

QY 75 KGWTCPCCKAGKVAHRADEISQVY--AH-QLGTCGSAVMTMTRHTAGORLHDLWTGLSA 131

A:Reference number: A91888; MUID:89008081; PMID:3170481
A:Accession: A31844
A:Molecule type: DNA
A:Residues: 1-456 <KEN>
A:Cross-references: UNIPROT:P22406
C:Genetics:
A:Genome: plasmid

Query Match 15.8%; Score 316; DB 2; Length 456;
Best Local Similarity 30.5%; Pred. No. 4.4e-18;
Matches 100; Conservative 40; Mismatches 112; Indels 76; Gaps 15;

QY 49 PISGV-----NGVTIVNGPKSGFGGLRSCGKMICPCCKAGKVAHRADEISQVVAQL 102
DB 3 PASGVIVAQAAGTSVVL-----GLMRCGRILWCPVCAATIRHKRAEIEITAAVVEVI 54

QY 103 GTGSAV-MVTMTMRTAQRLHDLWTGLSAWKAATNGRR-----WRTE----- 145
DB 55 KRGGTAYILVTFARHGHTDRLADLMDALQGTTRTPDPRPGAYQRLITGGTWAGRAKD 114

QY 146 -----REMYGCDGYRAVEITHGK-NGWHVHVHALLMFSGDV-----SENILESF 189
DB 115 GHRAADREGIRDIGVGMIRATEVTVGQINGWHPHIHAIVLVGKRTGERSAKQIVATF 174

QY 190 --SDAMFDRWTSKLVSLGFAAPLR-----NSGLDVRKIGIGHEAD-QVLAAYLTAKIA 237
DB 175 EPTGAALDEWQHWRSV-WTAALRKVNPAFTPDRHGVDFKRLETERDANDLAELYIAKQ 233

QY 238 SG--VCMFVSGDGKSGRGNAPWEI---AVDVG---DPOA-----LELWREPE 281
DB 234 DGKAPALEARADLTKATGNAVPELLOGLIGLGTGMDTDAAGVSGLENNLSRWHEYE 293

QY 282 FGSMDGRRALAWSGRLRARAGLGAELTDA 309
DB 294 RATRGRRALEWRYRLQMLGLDGDTEA 321

RESULT 3
S04020
hypoetical protein - Streptomyces sp. plasmid pSB24.2
C:Species: Streptomyces sp.
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: S04020
R:Polotin, A.P.; Sorokin, A.V.; Alekandrov, N.N.; Danilenko, V.N.; Kozlov, Y.I.
Dokl. Biochem. 283, 260-263, 1985
A:Title: Nucleotide sequence of DNA of the actinomycete plasmid pSB24.2.
A:Reference number: S04020
A:Accession: S04020
A:Molecule type: DNA
A:Residues: 1-437 <BOL>
A:Cross-references: UNIPROT:Q52205; EMBL:X03756
A:Note: the authors translated the codon TGG for residue 59 as Arg, CGG for residue 187
A:Note: the source is designated as Streptomyces cyanogenus
C:Genetics:
A:Genome: plasmid

Query Match 11.5%; Score 230.5; DB 2; Length 437;
Best Local Similarity 24.6%; Pred. No. 4.1e-11;
Matches 102; Conservative 43; Mismatches 140; Indels 129; Gaps 19;

QY 55 GVTIVNGPKSGFG-GLRSCGKMICPCCKAGKVAHRADEISQVVAQLGTGSA-MVTM 112
DB 6 GVTFARTAAAGAAVALGKRIWLCVPVCSQDPAPFSEITEAVVSLQGGWAYILVF 65

QY 113 TMRTAQRLHDLWTGL-----SAWKAATNGRRMTEREM 149
DB 66 TARTAAADRLDMDALQGTTRADTETGKRPAYQRLITGAAM--AGDKRKNQEGIR 123

QY 150 GCDGY---VRAVEITHGKN-NGWHVHVHALLMFSGDV-----SENILESFS---DAMP--- 194
DB 124 GRIGYIGMIRATEVTVGAGWHPHIHAIVLVGKRTGERSAKQIVATF 183

Db 69 KSLCLPLCNRRSGOSQNLQVLDEAHKORKTGRFLFLTLTAENASGENLKEVRKMG 128
 QY 132 AKKAATNGRRWRTEREMYGCDGYVRAVEITHGKNG-WHVHVALHMFSGDVSNILESPS 190
 Db 129 A-----ISKLFQYKKAENLLGVVRSSTEITINKNGYHQHMHVLLFVKPTFKDSANYIN 183
 QY 191 DAMFDRWTSKLSGLFAAPLNRSGGLDVKIGEGADQVLAAYLTKTIASGVGMEVSGDGK 250
 Db 184 QAEWSKLWQRAKMDY-QPVINVEAVRSNKAKGKNSLIASAQET----- 226
 QY 251 SGRHGNRAWEI-AVDAGGDQQA-LELWRETFEGSMGRRRAIAMSRL--RARAGLAGEL 306
 Db 227 -----AKYQVSKSDILNDQERDLQVVEDELEQGLAGSRQISYG-GLFKEIRKQLQLED 278
 QY 307 TDAQIVEQEE 316
 Db 279 VDAHLINVD 288

RESULT 6
 S25721
 replication protein repA - Streptococcus thermophilus
 C:Species: Streptococcus thermophilus
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S25721
 R.Janzén, T.; Kleinschmidt, J.; Neve, H.; Geis, A.
 FEMS Microbiol. Lett. 95, 175-180, 1992
 A:Title: Sequencing and characterization of pST1, a cryptic plasmid from Streptococcus
 A:Reference number: S25721
 A:Accession: S25721
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-315 <JAN>
 A:Cross-references: UNIPROT:Q56129; EMBL:X65856; NID:G47880; PID:G47881
 C:Superfamily: replication protein

Query Match 6.2%; Score 123.5; DB 2; Length 315;
 Best Local Similarity 22.8%; Pred. No. 0.016;
 Matches 50; Conservative 41; Mismatches 79; Indels 49; Gaps 11;

QY 12 KDRPVLVSSDKRGIRHELRLP-----KLOOI-----TTSETFNACGRPISGVNGVT 57
 Db 5 KKGQELVLDKNSRGKRDWRGRKILSLKLADIFKELQYKKTFFVERVISCGDTLQFIQND 64
 QY 58 IVNGPKSGFGGLR-----SCGKGWICPCAGKVGAAHRADEISQV-----AHQLGTGSVM 109
 Db 65 -----GNLKYQTYFC-KNKLCPMCNWRSMKYSYQTSRIVDEAIKQSPKGRFLF 113
 QY 110 VTMTMRHTAGORLHDLWTGLSAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNG--W 167
 Db 114 LTLTVKNVEGALNSTISQLTKSFDRLF--KRAKVQNRLL--GYLRSVEVTHNENKTY 168
 QY 168 HVHVALHMFSGDVSNILESPSDAM-----FDRWTSKL 201
 Db 169 HPHIHLVLM-----VRPSYFQSKDYITQKESDMSNSL 203

RESULT 7
 T44800
 replication protein repA [imported] - Streptococcus thermophilus plasmid pER35
 C:Species: Streptococcus thermophilus
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T44800
 R.Solow, B.T.; Somkuti, G.A.
 submitted to the EMBL Data Library, August 1999
 A:Description: Comparison of low molecular weight heat stress proteins encoded on plasmid
 A:Reference number: 222847
 A:Accession: T44800
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-315 <SL>
 A:Cross-references: UNIPROT:Q9RNW4; EMBL:AFI77167; PIDN:AAF04355.1
 A:Experimental source: strain ST135

C:Genetics:

A:Gene: repA
 A:Genome: plasmid pER35
 C:Superfamily: replication protein

Query Match

6.1%; Score 122.5; DB 2; Length 315;
 Best Local Similarity 21.6%; Pred. No. 0.019;
 Matches 48; Conservative 40; Mismatches 79; Indels 55; Gaps 9;

QY 12 KDRPVLVSSDKRGIRHELRLP-KLQOITTSFNACGRPISGVNGVTIYNGPKSGFGGL 70
 Db 5 KKGQELVLDKNSRGKRDWRGRKILSLKLADIFKELQYKKTFFVERVI----- 51
 QY 71 RSCG-----KGWICPCAGKVGAAHRADEISQV-----AHQLGTGS 106
 Db 52 -SCGDTLQFIQNDGTLKYQAYFCNKKLCPMCNWRSMKYSYQTSRIVDEAIKQSPKGR 110
 QY 107 VAMVTMRHTAGORLHDLWTGLSAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNG 166
 Db 111 FLTLTVKNVEGALNSTISQLTKSFDRLF--KRAKVQNRLL--GYLRSVEVTHNEND 165
 QY 167 --HVHVALHMFSGDVSNILESPSDAM-----FDRWTSKL 201
 Db 166 KTYHPIHLVLM-----VRPSYFQSKDYITQKESDMSNSL 203

RESULT 8

JN0856

reps protein - Streptococcus thermophilus (strain No.29) plasmid pST1

C:Species: Streptococcus thermophilus

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C:Accession: JN0856

R.Hashiba, H.; Takiguchi, R.; Joho, K.; Aoyama, K.; Hirota, T.

Biosci. Biotechnol. Biochem. 57, 1646-1649, 1993

A:Title: Identification of the replication region of Streptococcus thermophilus No.29

A:Reference number: JN0856; MUID:94080026; PMID:7764262

A:Accession: JN0856

A:Molecule type: DNA

A:Residues: 1-315 <HAS>

A:Cross-references: UNIPROT:Q9RNW8; UNIPROT:Q9RNW4

C:Genetics:

A:Genome: plasmid

C:Superfamily: replication protein

C:Keywords: replication initiation

Query Match

5.9%; Score 118.5; DB 2; Length 315;
 Best Local Similarity 22.4%; Pred. No. 0.04;
 Matches 52; Conservative 45; Mismatches 90; Indels 45; Gaps 12;

QY 12 KDRPVLVSSDKRGIRHELRLP-----KLOOI-----TTSETFNACGRPISGVNGVT 57
 Db 5 KKGQELVLDKNSRGKRDWRGRKILSLKLADIFKELQYKKTFFVERVISCGDTLQFIQND 64
 QY 58 IVNGPKSGFGGLR-----SCGKGWICPCAGKVG--AHRADEISQVVAHQLGTGSVM 109
 Db 65 -----GNLKYQAYFC-KNKLCPMCNWRSMKYSYQTSRIVDEAIKQSPKGRFLF 113
 QY 110 VTMTMRHTAGORLHDLWTGLSAWKAATNGRRWRTEREMYGCDGYVRAVEITHGKNG--W 167
 Db 114 LTLTVKNVEGALNSTISQLTKSFDRLF--KRAKVQNRLL--GYLRSVEVTHNENKTY 168
 QY 168 HVHVALHMFSGDVSNILESPSDAMF--DRWT--SKL--VSLGFAAPLNSG 214
 Db 169 HPHIHLVLMVRPSYFQSKDYITQAEWSDMWSLSKVDYVPMFIFIVKTEG 220

RESULT 9

A40651

replication protein A - Synecocystis sp. plasmid pCA2.4

C:Species: Synecocystis sp.

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C:Accession: A40651

R.Yang, X.; McFadden, B.A.

J. Bacteriol. 175, 3981-3991, 1993
A: Title: A small plasmid, pCA2.4, from the cyanobacterium *Synechocystis* sp. strain PCC 6803
A: Reference number: A40651; MUID: 93308076; PMID: 8320214
A: Accession: A40651
A: Status: preliminary
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-336 <YAN>
A: Cross-references: UNIPROT: Q06460; GB: I13739; NID: g408991; PIDN: AAA02970.1; PID: g290013
C: Genetics:
A: Genome: plasmid
C: Superfamily: replication protein

Query Match	5.7%	Score 115;	DB 2;	Length 336;
Best Local Similarity	26.2%;	Pred. No. 0.083;		
Matches	34;	Conservative 23;	Mismatches 45;	Indels 28; Gaps 6;

QY	79	CPCCACK---	VGAAHRADEISQVVAHQ	LGTSVAMVTWTRHTAGQRLHDLMTGLSAAWKA	135
Db	97	CPVCQWRRLMWRAKAFKILPQ	VEAYPKRFFELTLV	RNCGLQRLTITGMNQAWGR	156
QY	136	ATNGRWRTREMYGCDGYV	RAVEITHTGKNG-WhwHvHALLM-----	FS-GDVSENILES	188
Db	157	LVRKAW-----	PADGWIRSLVETIRKQGS	AHPHPCLLMWKASYFSGHYISQEV---	206
QY	189	FSDAMFDRWT	198		
Db	207	-----WT	208		

C:Species: Streptococcus thermophilus
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C:Accession: T46831
R:Somkuti, G.A.; Solaiman, D.K.Y.; Steinberg, D.H.
Plasmid 40, 61-72, 1998
A:Title: Structural and functional properties of the hsp16.4-bearing plasmid pER341 in *S. thermophilus*
A:Reference number: Z24102; MUID:98325270; PMID:9657935
A:Accession: T46831
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-314 <SOW>
A:Cross-references: UNIPROT:O30850; EMBL:AF019139; NID:g3342702; PIDN:AAC27655.1; PID:g3342702
A:Experimental source: strain 134
C:Genetics:
A:Gene: rep
A:Genome: plasmid pER341
C:Superfamily: replication protein

Query Match	5.7%;	Score 114;	DB 2;	Length 314;
Best Local Similarity	25.4%;	Pred. No. 0.093;		
Matches	47;	Conservative 33;	Mismatches 63;	Indels 42; Gaps 10;
QY	10	SGKDRPPVIVSDKGRHLEPPLKQQI-----TTSETFNACGRPSGVNGVTIVNG	61	: :
Db	17	NGKDR-----DWRG-RKLISLKLADIPELQYKKTFVERVISCGDLRFIQ-----	61	: :
QY	62	PKGSFGGLR-----SCGKGMTCPCCAGKVGAHRADEISQVV---AHQLGTGSAVMWTMT	113	: :
Db	62	---KQDGSILKLVQAYFC-KNKLCPICNNRRSMKYSQTSKIVDIAIKOEPKRFLFLTIT	117	: :
QY	114	MRHTAQRLHLDLWTGLSAAWKAATNGRAWTEREMYGCDYVRAVEITHGK--NCGWVHV	171	:::::
Db	118	VKNIEGKALNSTISQITSFDELFP--KRAKVORNLL---GYLSRVETIHENDNSVHPHI	172	:::::
QY	172	HALLM	176	: :
Db	173	HVLMM	177	: :

RESULT 11
A49783

	Matches	60; Conservative	28; Mismatches	112; Indels	54; Gaps	11;
QY	77	WICPCCAKVGAAH--RADEISQVVAHQLGCTGVSAMVTMTMRHTAGORLHDLMTGLSAAWK	134			
Db	127	WV-----AGEMDGLHLLRRSLSDAVVHPLEQGLAEP-----QELFLWILGERPAMW	174			
QY	135	AATNGR-----RWKTEREMYGCDGVYRAVEITHKNGWH-----VHVHALL	175			
Db	175	SAALERAALDEALRWGSPASFWGL--ALKAPATSAGNGTWMKPLFCAALALIVWVAGLN	232			
QY	176	MFGSDVSENILESDFAMDFRWTSKVLSIGFAA-----PLRNSGGLDVVKTKGGEADQVL	229			
Db	233	LYAQQLAE-----QGQSLQSQSSQRVQQAPELFWLVDPLFQARERRDAYLAGKADGTA	286			
QY	230	AAVLTKTIASGVGMVEGSDGKSGR---HGNRAPWEIANDVAGGDQFQALELWREFEFGSMG	286			
Db	287	APGLAALLHGAGEMPFLAGLQRLRDYHAGELDULELFLPGGDAAA---W-QGELGKHG	342			
QY	287	RRAIANSRGLRARA	300			
Db	343	LQADADSKGQVQRA	356			

RESULT 13
 140561
 replication initiation protein rep60 [validated] - Bacillus subtilis plasmids
 N/Contains: DNA ligase (EC 6.5.1.-); DNA lyase (EC 4.2.99.-)
 C/Species: Bacillus subtilis
 C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
 R/Accession: I40561; A32059; A43997; I40549
 R/Meijer, W.J.; Venema, G.; Bron, S.
 Nucleic Acids Res. 23, 612-619, 1995
 A/Title: Characterization of single strand origins of cryptic rolling-circle plasmids for
 A/Reference number: I40549; MUID:95206941; PMID:7899081
 A/Accession: I40561
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-340 <MEI>
 A/Cross-references: UNIPROT:Q45450; EMBL:U32380; NID:gl049123; PIDN:AAC4418.1; PID:gl04
 R/Devine, K.M.; Hogan, S.T.; Higgins, D.G.; McConnell, D.J.
 J. Bacteriol. 171, 1166-1172, 1989
 A/Title: Replication and segregational stability of Bacillus plasmid pBAAL.
 A/Reference number: A32059; MUID:89123116; PMID:2492507
 A/Accession: A32059
 A/Molecule type: DNA
 A/Residues: 34-124,'V',126-138,'E',140-230,'L',232-237,'E',239-340 <DEV>
 A/Cross-references: GB:M24251; NID:gl42426
 A/Experimental source: plasmid pBAAL
 A/Note: the authors' translation is shown for residues 313-333 and not the sequence show
 R/Darabi, A.; Forough, R.; Bhardwaj, G.; Watabe, M.; Goodarzi, G.; Gross, S.C.; Watabe,
 Plasmid 22, 281-286, 1989
 A/Title: Identification and nucleotide sequence of the minimal replicon of the low-copy-
 A/Reference number: A43997; MUID:90222329; PMID:2561213
 A/Accession: A43997
 A/Molecule type: DNA
 A/Residues: 34-124,'V',126-138,'E',140-230,'L',232-237,'E',239-301,'R' <DAR>
 A/Cross-references: GB:M34386; NID:gl44127; PIDN:AAA23002.1; PID:gl44129
 A/Experimental source: plasmid pBS2
 A/Note: the authors translated the codon ACG for residue 124 as Tyr, and GAG for residue
 C/Genetics:
 A/Gene: rep60
 A/Genome: Plasmid
 A/Start codon: ITG
 C/Superfamily: replication protein
 C/Keywords: carbon-oxygen lyase; ligase; plasmid replication
 F/252/Active site: Tyr (covalent DNA-binding) #status predicted

 Query Match 5.2%; Score 104.5; DB 2; Length 340;
 Best Local Similarity 24.2%; Pred. No. 0.61;
 Matches 38; Conservative 35; Mismatches 57; Indels 27; Gaps 9;

 98 ICPCGAKVG---AHRDEISQVVAHQIGTGSAVMTTMRHTAGQRHDLMTGLSAWK 134

Db 87 LCPNCARSLKTYAHNKLIVEANRQYGGWI--FLTLITRNVRKGLKPOISAMMEGFR 145
QY 135 AATNGRRWRTEREMYGCDGYVRAVEIT--HGKNGWHVHVHALL-MFSGDVSENTILESFD 191
Db 146 KLFQYKKVKT-----SVLGFRALEITKNHEEDYHPHFHVLIPVRKNYFGKNYIKQ--- 197
QY 192 AMPDRWTS---KLAVSLGFAAPLNSGGLDVRKIGGEA 225
Db 198 ---AETWLSWKAKWKLDY-TPI-----VDIRRVKGKA 225

RESULT 14
AG2059
heme biosynthesis protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AG2059
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriug
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2059
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-420 <Kur>
A:Cross-references: UNIPROT:Q8YVE9; GB:BA000019; PIDN:BA073728.1; PID:gl7131120; GSPDB
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all209

```
Query Match      5.2%; Score 104; DB 2; Length 420;
Best Local Similarity 24.4%; Pred. No. 0.86;
Matches 85; Conservative 47; Mismatches 135; Indels 82; Gaps 23;
```

Qy 35 QOITTSEFNACGRDIS-GVNGVTIVGPKSGSFGGLRS-----CGKMWICPCCAG 84
Db :|::||::||::||::||
Qy 36 QELSTAEALNLVQQAEVGIEKVTLIGGE-----AFLRPDWLVIKAITDAGMLCGMTTG 90
Db :|::||::||::||::||
Qy 85 KVGARADEISQVVAHQLGTSVAMVTMRHTAG-QLRLDLWTGLSAAWKAATNRRWR 143
Db :|::||::||::||::||
Qy 91 GYG-----ISLEMAQRMEAGIAKVSV--TDGWEATHDLRKGSKWSKSL----- 135
Db :|::||::||::||::||
Qy 144 TERMYGCDGVRAVEITHGNQHVVHALLMFSDGSVENILEFSDFAMDFTWTKL-V 202
Db :|::||::||::||::||
Qy 136 --RTM----SYLKVGILFGCN---QTINRL---SAPEFPSIYEHIRDAGARAWOIQLTV 183
Db :|::||::||::||::||
Qy 203 SLGFAP-----LRNSGLDVVRKIGGEADVLAAYLVTKIASCVGNESVDGKSGRHGNR 257
Db :|::||::||::||::||
Qy 184 PMGNAADNADILLQPSLLDD-----YPLMARVAQRANTEGVRVAAGN-NIGYYG-- 232
Db :|::||::||::||::||
Qy 258 APWHIAVDVGGDPQAELNREFBFG--SMGERRAIAWSRG---LFARAGLAGAELTD---A 309
Db :|::||::||::||::||
Qy 233 -PYE---RLLRGQGKEGWFRGCCAGLSLTLGLEADGTIKGCPSUPTTYAYAGNIERPLR 288
Db :|::||::||::||::||
Qy 310 QIVEQEESA PV-MVALTP---ARSWMIRTCAPVFVGHILGLVERAGATW 354
Db :|::||::||::||::||
Qy 289 EIVEQESELQNLYAGTEGYLWMGCKTK---FAE---LCRGGCTIW 331
Db :|::||::||::||::||

```

RESULT 15
AC3284
D-amino-acid dehydrogenase (EC 1.4.99.1) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 16-Aug-2004
C:Accession: AC3284
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Majer, C.; Los, T.; Ivanova,
.; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:title: The genome sequence of the facultative intracellular pathogen Brucella meliten
A:Reference number: AD3252; PMID:11756688

```

Sat Oct 23 17:44:54 2004

A;Accession: AC3284
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-420 <KUR>
A;Cross-references: UNIPROT:O8YJ30; GB:AE008917; PIDN:AAL51438.1; PID:GI7982146; GSPDB:G
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0256
A;Map position: I
C;Superfamily: Sarcosine oxidase
C;Keywords: oxidoreductase

Query Match 5.1%; Score 102.5; DB 2; Length 420;
Best Local Similarity 19.9%; Pred. No. 1.1;
Matches 63; Conservative 49; Mismatches 111; Indels 93; Gaps 12;
QY 95 SQVVAHQLGTGSVAVMTMRTAGORLHDLWTGLSAAMKAATNGRRWR--TEREMYGCD 152
DB 152 AELASQPGWDVREKAGIAYEHVGRRLAELQPLNPRFVAGTFVPGWKNVSDPKLF--- 208
QY 153 GYVRAVEITHGKNGWHVHVHALLMFSGDVSENILESFSDAMFDRWTSKLVSLGFAAPLNN 212
DB 209 -----GKAIW-----SYAESLGARFLSGKV---ASAKRQ 234
QY 213 SGGLDVRKIGGEADQVLAAYLTAKIASCVGMEVSGDGKSGRHNRAPEIAVDVGGDPQ 272
DB 235 NGGVRLRLENG--EINATHVLVLMAGAWSRDLAKGFG-----DIVELDT 277
QY 273 -----ALELWREFEFGSMGRRAIAWSRGLRARAGL---GAEL-----TDAQIVE 313
DB 278 RGYNTLPVGSFVVKQLTFPGHGVTTPMETGLRVGGAVEFGGLDLPNFAESEM LKK 337
QY 314 QEESAPVMVAIIPARSMVMIRTCA-----YVFG-ILGLVEAGATWEN 356
DB 338 ASKFLFGL-KVEGGRQWGYRPSMPDPLPVIGRASAGNVYGFHGHGLGTQSAATARL 396
QY 357 LRDLHLRYLPADVRP 372
DB 397 IRDLITGSEFAIDIEP 412

Search completed: October 23, 2004, 03:58:09
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2004, 04:02:18 ; Search time 3866 Seconds
(without alignments)
3572.338 Million cell updates/sec

Title: US-10-007-527A-2
Perfect score: 2005
Sequence: 1 MTSVSAEHLGKDRPVLVS.....HLHRLPAADVRRPPIISVRK 379

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_spool/US10007527/runat_22102004_114116_26629/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10007527@cgn 1.1 3437 @runat_22102004_114116_26629 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsl1:*
9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	129.5	6.5	929	5	BUS19534
C 2	129	6.4	1648	4	BM913094 AGENCOURT
C 3	128.5	6.4	1090	4	BM457396 AGENCOURT
C 4	127	6.3	2033	9	AG347460 Mus muscu
C 5	122.5	6.1	894	4	B1661991
C 6	122	6.1	2887	9	AV411933 Homo sapi
C 7	120	6.0	918	9	CC623500
C 8	120	6.0	1255	9	CL479652
C 9	119.5	6.0	1661	8	CC321442

C 10	119	5.9	950	5	BUS19534
C 11	119	5.9	1430	4	BM907785
C 12	119	5.9	1687	4	BI224374
C 13	118	5.9	851	5	BQ714786
C 14	118	5.9	893	8	BZ554247
C 15	117.5	5.9	1011	5	BX898881
C 16	117	5.8	1306	5	BQ936859
C 17	117	5.8	1615	2	BF783863
C 18	116.5	5.8	689	5	BUT05875
C 19	116.5	5.8	695	7	CF730403
C 20	116.5	5.8	697	7	CF724491
C 21	116	5.8	623	7	CF433522
C 22	116	5.8	867	4	BI411296
C 23	116	5.8	943	6	CA280312
C 24	116	5.8	1199	4	BG684181
C 25	115.5	5.8	703	6	CB248284
C 26	115.5	5.8	786	5	BQ572893
C 27	115.5	5.8	799	6	CA320693
C 28	115.5	5.8	857	6	CAV50314
C 29	115.5	5.8	918	6	BY717387
C 30	115.5	5.8	2161	3	AK049163
C 31	115.5	5.8	2161	3	AK083015
C 32	115.5	5.8	2162	3	AK028287
C 33	115.5	5.8	2167	3	BC013842
C 34	115.5	5.8	2168	3	AK017609
C 35	115	5.7	696	6	BY729572
C 36	115	5.7	899	8	BZ551248
C 37	114.5	5.7	745	9	CL829504
C 38	114.5	5.7	930	4	BG922360
C 39	114.5	5.7	1189	8	CC318464
C 40	114	5.7	787	7	CF878630
C 41	114	5.7	990	7	CF878672
C 42	114	5.7	1937	9	AG279610
C 43	113.5	5.7	974	5	BQ672428
C 44	113.5	5.7	1349	9	CL495997
C 45	113.5	5.7	1414	5	BQ276790

ALIGNMENTS

RESULT 1
BUS19534/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BUS19534 929 bp mRNA linear EST 12-SEP-2002
AGENCOURT_10166703 NIH_MGC_134 Mus musculus cDNA clone
IMAGE:6517370 5', mRNA sequence.
BUS19534
BUS19534.1 GI:22827060
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
1 (bases 1 to 929)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILLNL at:
<http://image.llnl.gov>
Plate: LLAM4097 row: g column: 03
High quality sequence start: 6
High quality sequence stop: 710.
Location/Qualifiers
1..929
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

FEATURES
source

/clone="IMAGE:6517370"
/tissue type="undifferentiated limb"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH MGC 134"
/note="Vector: PCMV-SPORE6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 0.0607 Length: 929
Score: 129.50 Matches: 91
Percent Similarity: 31.90% Conservativity: 28
Best Local Similarity: 24.40% Mismatches: 95
Query Match: 6.46% Indels: 160
DB: 5 Gaps: 19

US-10-007-527A-2 (1-379) x BU519534 (1-929)

QY 20 SerSerAspLysArgGlyLeuArgHisGluLeuArgProLys-----Leu 34
DB 768 GCTTCTCAAAAGCGCTTGGGCATCATCTT---CCTGAACACCCCTTGTGACTTTG 712
QY 35 GlnGlnIleThrSerGluThrPhe-AsnAlaCysGlyArgProLysSerGlyValAs 54
DB 711 CAAAAGGTACACAGTCTCAGAGCGATTGTGGAGCATGCTCACAA----- 669
QY 54 nGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyLeuArgSerCysG 74
DB 669 ----- 669
QY 74 YLysGlyTrpIleCysProCysAlaGlyLysValGlyAlaHisArgAlaAspGlu 94
DB 668 -----TGG---TGTCATGCTGC----- 654
QY 94 eSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThr 114
DB 653 -----GGAACATCTCACTAGTACCTGATCTGTAGGGT 619
QY 114 tArgHisThrAlaGlyGlnArgLeuHisAspLeuThrGlyLeuSerAlaAlaTrpLy 134
DB 618 A-----GGACAGGCTGGAGACCA-----GGCTGAGCCAGACATGGAA 578
QY 134 sAlaAlaThrAsnGlyArgArgTrp-ArgThrGlu-----ArgGluMetTy-Gly-Cys 151
DB 577 G-----CGCTGTTAGGACGAGGCTTCTCCGACACATGTTGCTCTGT 536
QY 152 AspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGly---TrpHisValHis 170
DB 535 -----GGGCACCTTGGCATCTGGGTACTAGG 509
QY 171 ValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPheSer 190
DB 508 CTCACAGCATCTCTGTGACTCGGC----- 482
QY 191 AspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaProLeu 210
DB 481 -----TCTCAGGGCCCGTGCCTGGGCTCCCACTTCCCTGC 446
QY 211 ArgAsnSerGlyGlyLeuAspValArgLysIleGlyGlyGluAlaAspGlnValLeuAla 230
DB 445 CCGCCTCAGCAGACATCTGTGAGCTCAGA-----GACGGAGTCTCTGA 404
QY 231 AlaTyrlLeuThrLysIleAlaSerGlyVal---GlyMetGluValGlySerGlyAspGly 249
DB 403 GGGGATCTGCGCGCTGTTGACCAACATTAACAGTTGTACTCTTGTTCGTTGACGGG 344
QY 250 LysSerGlyArgHisGlnArgAlaProTrpGluLeuAlaValAspAlaValGlyGly 269
DB 343 GGATGGGCAACCGGGGAGCCAGGCTCAGCTGGAGGGCTAGGGTCAGC----- 296
QY 270 AspProGlnAlaLeuGluLeuTrpArgGluPheGluPheGlySerMetGlyArgAla 289

Db 295 -----TGG-----AGGCTGGG 284
QY 290 IleAlaTrpSerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAla 309
Db 283 CTCAGCTGGGAACTGGGCTCAGGAGAACTGGGGGAGG----- 242
QY 310 GlnIleValGluGlnGluSerAlaProValMetValAlaIleIleProAlaArgSer 329
Db 241 -----GTCCAATCCAGCAGCCGTAACT-----CCGGCGCGC--- 209
QY 330 TrpMetMetIleArgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGlu 349
Db 208 -----GAGCGCGCGGGCTGGTCACT 188
QY 350 AlaGlyAlaThrTrp-GluAsnLeuArgAspHisLeuHisTyArgLeuProAlaAlaAs 369
Db 187 CCGGGTTCGGCTGGGGCGCGCTTGGGCTGTCTCCGGTGACACGACGAGGC 128
QY 369 pValArgProProlIleSerValArgLys 379
Db 127 CACGAGGCCACCGAGGCTCGGGGTGAGACAG 97

RESULT 2
BM913094/c
LOCUS
DEFINITION
AGENCOURT 6613820 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5475102
5', mRNA sequence.
BM913094
VERSION
BM913094.1 GI:19363473
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1648)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1989 row: g column: 07
High quality sequence stop: 263.

FEATURES
source
1..1648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5475102"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: Skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

Alignment Scores:
Pred. No.: 0.151 Length: 1648
Score: 129.00 Matches: 93
Percent Similarity: 30.73% Conservativity: 33

```

Best Local Similarity: 22.68%      Mismatches: 153
Query Match: 6.43%      Indels: 131
DB: 4      Gaps: 17

US-10-007-527A-2 (1-379) x BM913094 (1-1648)

QY 10 SerGlyLysAspArgProValLeuValSerSerAspLysArgGlyLeuArgHisGlu 29
   |||||
   |||||
   |||||
Db 1101 GCTGCCAGAGCCGCCACATCAATATCAC-----CCACTGGGGTCCGACACCA 1048
   |||||
QY 30 LeuArgProLys-----
   |||||
Db 1047 AAGCGTCCAGCGGTCCGCCAGCAGGAGCGTGTGACTGTCTCTTACACAGATTG 988
   |||||
QY 34 ---LeuGlnGlnIleThrThrSerGluThrPheAsnAlaCysGlyArgPro----- 49
   |||||
Db 987 CCGTTGAGGCACTCGCCAGGNAG-----GGCGCGGTGAGACTGTA 946
   |||||
QY 50 -----IleSerGlyValAsn-----
   |||||
Db 945 ATGTTGCTAGAGGGCGTTAATACTCTGCGTGGCTGATGTATGTGACGAGGAGGACG 886
   |||||
QY 55 -----GlyValThrIleValAsnGlyProLysGly 64
   |||||
Db 885 GTGTTCAATTGTCGTCGCGCGCTTAGTAGTGCACAGAGTGTGTCTCGCACCATGTGT 826
   |||||
QY 65 SerGlyPheGlyGlyLeu-ArgSerCys-----GlyLysGlyT 77
   |||||
Db 825 GCGGTCGTCTGGGCTCACGTGCTGCTGGGCTGGGAGTGCCTGCTGGGCGTGGAT 766
   |||||
QY 77 rPileCysProCys-----CysAlaGlyLysValGlyAlaHisArgA 91
   |||||
Db 765 GGGTGTGTTTGGCGCATTTGGAGCGCGCTGTCGCCGATCTGTGCTGCTCGCGCG 706
   |||||
QY 91 laAspGluIleSerGlnValValAlaHisGlnLeuGly---ThrGlySerValAlaMetV 110
   |||||
Db 705 CATCTGTTTTTGGCTGCTCGCGCACTAGCATGCGCGCGCGCAAGGTCTCTCATCGCA 646
   |||||
QY 110 alThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTrpThrGlyLeus 130
   |||||
Db 645 TTCAGCCGAGCGCGGCATCCGCC-----GAGGGGCGCG 610
   |||||
QY 130 erAlaAlaTrpLysAlaAlaThrAsnGlyArgArgTrpArgGluArgGluMetTyrG 150
   |||||
Db 609 GGGGATCTGG---GTTGCGCGCCAGCGCATGTGCGGCTTCCCGGCGACAGCTCG 553
   |||||
QY 150 lCysAspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGly-TrpHisVal 169
   |||||
Db 552 GGGCGCTGGC-----GGTAGTCGCGGTTTGGCACTCC 520
   |||||
QY 170 HisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPhe 189
   |||||
Db 519 GAGACGGGGGATGGTGTGACTGTGGG----- 490
   |||||
QY 190 SerAspAlaMetPheAspArgTrpThrSerLysLeu----- 201
   |||||
Db 489 -----TGGACGGCGCTCAGCAGTTCGTGGCGGGCTCGCCCGT 451
   |||||
QY 202 ---ValSerLeuGly-PheAlaAlaProLeuArgAsnSerGlyGlyLeuAspValArgLy 220
   |||||
Db 450 GTGTGGCGGTGGCGTTTCCAAACGGCGCGCGCTATAGTGCACGCTTGGCGCTTTC 391
   |||||
QY 220 sIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyVa 240
   |||||
Db 390 TCGTGGCGGTTCGTTTGAAGTTCGCGGGGCTCGTCACACAGTTTGTGCG----- 336
   |||||
QY 240 lGlyMetGluValGlySerGlyAspGlyLysSerGlyArgHisGly-----As 256
   |||||
Db 335 -GACATGTGCGGGCGCGCGGCTTGTGCGGGAGCGCGGGTCCCGGACCGAGAC 277
   |||||
QY 256 nArgAlaProTrpGluIleAlaValAlaAspAlaValGlyAspProGlnAlaLeuGluLe 276
   |||||
Db 276 GGAGGGGCTTGG-----ACGCGGAGGGTGTGCGCGCGCGACCTGGCGCG 229

```

```

QY 276 uTrpArgGluPheGlu-PheGlySerMetGlyArgAlaIleAlaTrp----- 292
   |||||
Db 228 GTGGCGCGCGCGCGCGACAGCGCGCGCGCTCGCAAGCTGGTCTCTGGAA 169
   |||||
QY 293 -----SerArgGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnI 311
   |||||
Db 168 GAGGCGCGAGAGCGCGCGGGTGGGTCTGGAGCGGAGACCCGCGAGGCG 109
   |||||
QY 311 leValGluGlnGluSerAlaProValMetValAlaIleIleProAlaArgSerTrpM 331
   |||||
Db 108 TCGCGCGCTCGGAGGCGCGGCGAAGCGCGTGTGAGCGGTGCGCGCGCTCGCGCG 49
   |||||
QY 331 etMetIleArgThrCysAla 337
   |||||
Db 48 CGTAGCTCGATCCCGGCT 29

RESULT 3
BM457396 1090 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6407558 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5584861
DEFINITION 5', mRNA sequence.
ACCESSION BM457396
VERSION BM457396.1 GI:18506436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1090)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12350 row: d column: 14
High quality sequence stop: 538.
FEATURES
    source
        1..1090
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5584861"
            /tissue_type="embryonal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_92"
            /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 2.5 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 0.0949 Length: 1090
Score: 128.50 Matches: 80
Percent Similarity: 31.38% Conservative: 27
Best Local Similarity: 23.46% Mismatches: 94
Query Match: 6.41% Indels: 140
DB: 4 Gaps: 15

US-10-007-527A-2 (1-379) x BM457396 (1-1090)

QY 7 GluHisLeu-----SerGlyLysAspArgProValLeu----- 18
   |||||
Db 314 GAGCACTGACGACTAGCGCCAGAGGTGCCCGCGTCTAAAGAGCTGTGCAGATTT 373

```



```

Db 1716 GGAGTCGGGTGTTGAGTCAGGTGGTGGTGGCTTGGTGTGGCGGCTGGCGGGTGT 1657
QY 87 GlyAla----- 88
Db 1656 GCGGCGCGCTCGAGCGCGCTTGGCGGGGCGTGGCGCGCGCGGGTGTGCGG 1597
QY 89 -----HisArgAlaAspGluIleSerGlnValValAlaHisGln 101
Db 1596 GCGGCGTGGTGGCGCGGTCACAGCCCTAGGCGATGCGGTTGGGGTGTGCGCGGT 1537
QY 102 LeuGlyThrGly----- 105
Db 1536 CTGGGTGCGCGCTGCGTGGTGGTGGCGGTCTTGGCTCTCTCAGGTGGTATT 1477
QY 106 -----SerValAlaMetValThrMetThrMetArg 115
Db 1476 GGATTCCGCGTGGCGGTGCTGAGCGTGGCGGAGTGGTGTGGCGGTGCGGCGAGA 1417
QY 116 HisThrAlaGlyGlnArgLeuHisAspLeuTrp----- 126
Db 1416 GCGTCGCGCGGAGCGTGTATCGCGCGGTGGTGGTAGTTCGCGCGCGGTGATGGG 1357
QY 127 -----ThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArg 140
Db 1356 CTTGAGGCTGCGGGGGTGGCGGTGGCGGTGGCATGGATGCAGCGCGCGAGTTCA 1297
QY 141 ArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluLe 160
Db 1296 GCGTCTCGTAGCCCGG-----GTGAGACTG 1270
QY 161 ThrHisGlyLysAsnGlyTrpHisValHisValHisAlaLeuLeuMet----- 176
Db 1269 CTGGGGGGGTGTGGATGGAGGTCGCGTGGCGGTGGCGGCTGGCGCGCCAGCGG 1210
QY 177 -----PheSerGlyAsp----- 180
Db 1209 GCGGGGGTGGCGGAGGTGGTGGTGGGAGGCTTCGTCGGCGCGCGAGCGGGTG 1150
QY 181 ---ValSerGluAsnIleLeuGluSerPheSerAlaMetPheAspArgTrpThrSer 199
Db 1149 GCGGTAGGCGCGGGGTGCTG-----TTGTGGAGCGG 1117
QY 200 LysLeuValSerLeu-----GlyPheAlaAlaProLeuArgAsnSerGlyGlyLeu 216
Db 1116 CGGCTCGTCGCGTGGGGGTGGCGGTCGCGGTGGTGGTGGTGGTGGTGGTGGTGG 1057
QY 217 AspValArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIle 236
Db 1056 -----GCGGGGGGGGTGGCGCGTGGCGGTGGCGCGCGCGCGCGCGCGG 1009
QY 237 AlaSerGlyValGlyMetGlu-----ValGlySerGlyAspGly 249
Db 1008 TCGGGGGCGTGGAGTGTCTCCATGCCGCTGGGCGCGGTGTGTGTGTGTGTGTGTCA 949
QY 250 LysSerGlyArgHisGlyAsnArgAlaProTrpGluIleAlaValAspAlaValGlyGly 269
Db 948 GCTGGGGGTGGGGGGTGTCTCGCGCGGTGGTGGGGGCGGTGGGGGCGGCGGT 889
QY 270 AspProGlnAlaLeuGluLeuTrpArgGluPheGluPheGlySerMetGlyArgArgAla 289
Db 888 GGGCCC---GCATGTGGCGGTGG-----GGTGGCGCGCCCTTCGGGGCC 847
QY 290 IleAlaTrp-----SerArgGlyLeuArgAlaAlaGlyLeuGly 303
Db 846 GGGGGGTGGCGGGGTGGGGCGCTCTGCGCGCGGGTGTGGCGCTCCCGGGGTGGGT 787

```

RESULT 5

```

BI661991/c
LOCUS 603304986F1 NCI_COAP_Mam4 894 bp mRNA linear EST 12-SEP-2001
DEFINITION 603304986F1 NCI_COAP_Mam4 Mus musculus cDNA clone IMAGE:5350797 5',
ACCESSION BI661991
VERSION BI661991.1 GI:15576227

```

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 894)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM1891 row: k column: 22

High quality sequence stop: 849.

Location/Qualifiers

FEATURES

source

1..894

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NMRI"

/db_xref="taxon:10090"

/clone="IMAGE:5350797"

/tissue_type="tumor, gross tissue"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI_COAP_Mam4"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

ORIGIN

Alignment Scores:

```

Pred. No.:      0.276      Length:      894
Score:          122.50     Matches:    79
Percent Similarity: 30.77%  Conservaive: 25
Best Local Similarity: 23.37% Mismatches:   114
Query Match:     6.11%     Indels:     120
DB:              4        Gaps:       16

```

US-10-007-527A-2 (1-379) x BI661991 (1-894)

```

QY 68 GlyGlyLeuArgSerCysGlyLysGly-----Trp-----IleCysPro 80
Db 811 GGGACATTGCATCTCTCGGGAAGCATACGACGTGGGATACACGCTTTTGATCTTGCCCA 752
QY 81 CysCysAla-----GlyLysValGlyAlaHisArgAlaAspGluLeuSer 95
Db 751 CATTGCGCTCTTCAAAGCGCTTGGCAFCATCTCTCTGGACACCTTGTCTGAC----- 701
QY 96 GlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetValThrMetThrMetArg 115
Db 700 -----TTTGGCAAGGTCTACAGTCTCAGAGCGATTGTGGAGCATGCT 659
QY 116 HisThr-----AlaGlyGlnArgLeu-----HisAspLeuTrpThrGlyLeuSer 130
Db 658 CACAATGTGTCCATGCTGCGGAACATCTCACTAGCACCTGATACCTTTAGGTAGGATAGGAC 599
QY 131 AlaAlaTrpLysAla-----AlaThrAsnGlyArgArgTrpArg 143
Db 598 AAGGCTCTGGGAGCCAGGCTGAGCGATGGAAGCGCTGGTAGGCGAGGCTTTCTC 539
QY 144 ThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGly 163
Db 538 GACATGTTGCTCTGTGGGCACCTTGGCATC----- 506

```


Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miquel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwalis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS801460; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not single contiguous sequences.
Class: TDNA tagged.

FEATURES

source
1. .1255
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL 30 C09.v1"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN

Alignment Scores:
Pred. No.: 0.773 Length: 1255
Score: 120.00 Matches: 94
Percent Similarity: 30.00% Conservative: 26
Best Local Similarity: 23.50% Mismatches: 106
Query Match: 5.99% Indels: 176
DB: 9 Gaps: 21

US-10-007-527A-2 (1-379) x CL479652 (1-1255)

QY 47 GlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySerGly 66
DB 1154 GGGCGGGGATTCGGGGTGTC-----CTGGTGGGGGCTTTAGAGGCAGGGGG 1107
QY 67 PheGlyGlyLeuArgSerCysGlyLysGlyTTP-----IleCysProCys--- 81
DB 1106 TGGTGTAGCGTGGGCGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1047
QY 82 CysAlaGly-----LysValGly----- 87
DB 1046 TGTGCGGTGGCGCGCGCGGTCTGGGCGCGCGGTGGGCGATGATGCGTAGGTGGGG 987
QY 88 -----AlaHisArgAlaAspGluIleSerGlnValValAlaHisGlnLeuGlyThr 104
DB 986 CGCGNAGCGTTCGATNGTCGGGAACGTCGCGCGCATGCCGGGACGCCAGCA-GGTAGC 928
QY 105 GlySerValAlaMetValThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisasp 124
DB 927 GGGGGTGA-----AGGAGTCCGTAGTCTCGACTA----- 895
QY 125 LeuTrpThrGlyLeuSerAlaAlaTrpLysAlaAlaThrAsnGlyArgArg-----Trp 142
DB 894 -----GCTGTCTGTAAAGTGTGGTGTG-TGTGGAGCATNGGGCGGGAATTGTCGG 842
QY 143 ArgThrGluArgGluMetTyGlyCysAspGlyThrValArgAlaValGluIleThrHis 162
DB 841 CCGTGTGTGGCGATG-----GGACAAGGTCCGGTAGTCAGCGCGCAGGAG 797
QY 163 GlyLysAsnGly-----TrpHisValHisValHisAlaLeuLeuMetPheSer 178
DB 796 GGGCAAGGTGGCAGTGGGACAGATGGGTGTGACGGGTAGTGTGTCGTG----- 746
QY 179 GlyAspValSerGluAsnIleLeuGluSerPheSerAspAlaMetPheAspArgTrpThr 198

DB 746 ----- 746
QY 199 SerLysLeuValSerLeuGlyPheAlaPro-----LeuArgAsnSerGlyGlyLeu 216
DB 745 GGGCGGTGCGCGCGTCCGGTCCGGAGAGCCCTCGGGCGTCCAGGAGCGGGGT--- 689
QY 217 AspVal-ArgLysIleGlyGlyGluAlaAspGlnValLeuAlaAlaTyLeuThrLysI 236
DB 688 ---GTGTAGGTGAGTGGGGGAGCGG----- 664
QY 236 eAlaSerGlyValGlyMetGluValGlySerGlyAspGlyLysSerGlyArg----- 253
DB 663 -CGGGGGGGCGGTGGGTAAAGGGGGGGTGGGGGGCGACGCACTGTGTGGAGGGAGAGGG 605
QY 254 -----HisGlyAsnArgAlaPro----- 259
DB 604 CGTGGCGGGGGGAGCGGGAGCGGGGGGGGGGAGCGGGGAGCGGGGAGCGGAGGG 545
QY 260 -----TrpGluIleAlaVal----- 264
DB 544 GGGGGGGGGGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 485
QY 265 -----AspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheG 281
DB 484 GGGGGGGGGGGTGGGGGGGGTGGTCTTCTGTGGTGGTGGGGGGTGGGG----- 433
QY 281 uPheGlySerMetGlyArgArgAlaIleAlaTrp-----SerAr 294
DB 432 ---GGGGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 377
QY 294 gGlyLeuArgAlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluCl 314
DB 376 GGGCGGTGCGGGCGCGGGGGTGGTGGGG-----GGGTGGGGCGCGTTCG 320
QY 314 nGluGluSerAlaProValMetValAlaIleProAlaArgSerTrpMetMetIleAr 334
DB 348 -----CG 347
QY 334 gThrCysAlaProTyValPheGlyGluIleLeuGlyLeuValGluAlaGlyAlaThrTr 354
DB 346 TTCGTGCGCG-----GGGTGGGGCGCGTTCG 320
QY 354 pGluAsnLeuArgAspHisLeuHisTyArgLeuProAlaAlaAspValArgProPro 373
DB 319 CGGGGGTGTGGTGGGACCCCGCGCGCGTGGGGGGGGGGGCTTCCCTCCGCCCC 262
RESULT 9
CC321442 1661 bp DNA linear GSS 14-MAY-2003
LOCUS TAM32-34D5_ECL1.1 TAM32 Gallus gallus genomic clone TAM32-34D5,
DEFINITION genomic survey sequence.
ACCESSION CC321442
VERSION CC321442.1 GI:30715500
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1661)
AUTHORS Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 152000 Std Error: 0.00
Seq primer: EC1 TACGACTCACTATAGGGCG
Class: BAC ends
High quality sequence start: 74

ORIGIN

Alignment Scores:
Pred. No.: 0.657 Length: 950
Score: 119.00 Matches: 80
Percent Similarity: 30.08% Conservative: 31
Best Local Similarity: 21.68% Mismatches: 120
Query Match: 5.94% Indels: 139
DB: 5 Gaps: 17

US-10-007-527A-2 (1-379) x BU2711357 (1-950)

QY 43 PheAsnAlaCysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyPro 62
DB 909 TAAGTCTTCCGCGGTGTGTGGGGTGTCCGGGGTGTCCAGTGGCCCTCGGTTG 850
QY 63 LysGlySerGlyPheGly-----GlyLeuArgSerCysGlyLysGlyTyr 77
DB 849 TTGGGGGTGGGTGTGGTGTGTCTTTTGGGGGGCGAGGGGTGCCATGGGGATGG 790
QY 78 IleCysProCysCysAlaGlyLysValGlyAlaHisArgAlaAspGluIleSerGlnVal 97
DB 789 GTC-----TGCTGTGCTCGCGGTGTGGG----- 766
QY 98 ValAlaHisGlnLeuGlyThrGlySerVal-AlaMetValThrMetThrMetArgHisTh 117
DB 765 -----CTTGGGATGGGCCCTTGTGCTGCGGTGTTGG----- 732
QY 117 xAlaGlyGlnArgLeuHisAspLeuThrGlyLeuSerAlaAlaTyr-LysAlaAlaT 137
DB 731 -----GTTTGGTGTGGT-----TGTTGGTTTGGCGGTGG 703
QY 137 hrAsnGly-----ArgArgTyrArgThrGluArgGluMetTyrGlyCysAspG 153
DB 702 CAGGGGGTGTGTGTGCTTGTGGGTGGGGCGTGCATGCGGTGTGGTGGCGGTGGGG 643
QY 153 lYtyrValArgAlaValGluIleThrHisGlyLysAsnGly-----T 167
DB 642 GCAGGGTGTGTGTGCTTCTTCTATCATGCGGGCGGGGCTGCTTCTGGGCTGAGGTGT 593
QY 167 rPhisValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuG 187
DB 582 GG-----GGGGAGGTTCGCCAGGG----- 563
QY 187 luSerPheSerAspAlaMetPheAspArgTyrThrSerLysLeuValSerLeuGlyPheA 207
DB 562 -----TGTTGGGGGCTGG 550
QY 207 laAlaProLeuArgAsnSerGlyLeuAspValArgLysIleGlyGluAlaAspG 227
DB 549 TTTGAGGCTGCGTGGTGTGTGGTCCGGGTTCGGGTGTTGCGAGGTGTGTGCTGGG 490
QY 227 lnVal-----LeuAlaAlaTyrLeuThrLysIleAlaSerGlyValG 241
DB 489 CCGTGAGTTCAGTGGTGGGGCCATCTCAGGTGAGCTTAGTTGTGCGAGCCTGGGATCAG 430
QY 241 lYMetGluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrog 261
DB 429 GCGTGGAGCAGGGGTGTGGA-----TGCG 406
QY 261 luIleAlaValAspAlaValAlGlyLysProGlnAlaLeuGluLeuTyrArgGluPheG 281
DB 405 GAGTGAGGTAGGAGGCAAGTCTGGGCAA----- 377
QY 281 luPheGlySerMetGlyArgArgAlaIleAlaTyr-----SerArgGlyLeuArgAlaA 299
DB 376 -----GGCCCCGAGGCAAGGCCGTAGGCTGTGGTATCGTTCGATGGGGATCGAGCA 322
QY 299 rGAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerAlaAlap 319
DB 321 GAGCA-----TCAGCGC 310
QY 319 roValMetValAlaIleIleProAlaArg-----SerTyrMetMetIleA 334

Db 309 CGTGGGCTCAGAGTTTGTCTTCCAAACGACGAGGTTCAGATGAAGCTTCATGGCTTCAGGCAG 250
QY 334 rgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeuValGlu----- 349
Db 249 CAAAGAAGCTGAGTTGTCAGTTTCAGGTCAGGAGATCTGTGGGGCCACAGCAACACAGGACA 190
QY 350 --AlaGlyAlaThrTrpGluAsnLeuArgAspHisLeuHisTyrArg-----LeuP 366
Db 189 GGGCTGGAAACCATCAGATAGCATTCCTTTAGCCAGCATTTAAAGCCCTGCTGCTCC 130
QY 366 roAlaAlaAspValArgProPro 373
Db 129 CTGCAGCATG-GAGCGTCCGCCT 108

RESULT 11

BM907785/c

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

BM907785

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1430)

NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: DCTD/DP/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12744 row: f column: 10

High quality sequence stop: 319.

FEATURES

source

1. .1430

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5736201"

/tissue_type="large cell carcinoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 68"

/notes="Organ: lung; Vector: pCMV-SPORT6; Site: 1; NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.8 kb. Library constructed by Life

Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 1.16 Length: 1430

Score: 119.00 Matches: 90

Percent Similarity: 34.82% Conservative: 43

Best Local Similarity: 23.56% Mismatches: 110

Query Match: 5.94% Indels: 139

DB: 4 Gaps: 21

US-10-007-527A-2 (1-379) x BM907785 (1-1430)

QY 46 CysGlyArgProIleSerGlyValAsnGlyValThrIleValAsnGlyProLysGlySer 65

Db 1343 TGTGGGGGCCCA-----GGTATGGGGGGGTATTGG-----GGGGGTGGGGGAAA 1299

QY 66 GlyPheGlyGlyLeuArgSerCysGlyLys----- 75

```
Db 1298 GGTGTGGGTGATTTTAAATGTTGGTCCGTCCTCCGTTATGTTAGGGGGTGGCGGGGG 1239
Qy 76 -----GlyTrrpIleCys 79
Db 1238 GTGGATATGAGAGTGGGGGTGACGTGCCCTCATATGACAGCGGTGGGTAATTTTGT 1179
Qy 80 ProCys-----CysAlaGlyLysValGlyAlaHisArgAlaAsp 92
Db 1178 CAGTGTGGCGAGGATATGGGTATTTATACGGGGGGCGATTCCGA-----AGTGGAGN 1125
Qy 93 GluIleSerGlnValAlaHisGlnLeuGly-ThrGlySer-----ValAlaMetVa 110
Db 1124 NGCTCTCTAGTCTGTATGCGGTATATCGCCATAGCGCGTTCATGAATAGAGCTCGT 1065
Qy 110 lThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuThrThrGlyLeu-- 129
Db 1064 GACTATCTGCT---AAACAT-----TTGTGGGTGGTGTGCTG 1032
Qy 130 -----SerAlaAlaTrrp-----Lys-AlaAlaThrAsnG 139
Db 1031 CGGGACCGGGCGAGTCAGTGGCGCTTTAGTGGGTTTTGTCGGCGTGGGGATAACGCTG 972
Qy 139 lYArgArgTrpArgThrGluArgGluMetTyrGlyCysAspGlyTyrValArgAlaValG 159
Db 971 GTAGCCGCGCGGTGAGAGAGCGAGTAAATGTGACGG-----GCTTTAT 921
Qy 159 luIleThrHisGlyLysAsnGlyTrrpHisValHisValHisAlaLeuLeuMetPheSerG 179
Db 920 GTCTAGACTAGGCGCTTAGAGGGGTGGCAGGGTGTGTACGAGCTTTTGTGTGTGG 861
Qy 179 lYAspVal-----SerGluAsnIleLeuGluSerPheS 190
Db 860 GGGCGGTAGGATAATGTGCGCGGTAAAGAGAGTCCAGAAATGGTGTGATCGGCATAT 801
Qy 190 erAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAProL 210
Db 800 CGAACAGTCAT-----GTCAGCCCTATATATGCG-----G 771
Qy 210 euArgAsnSerGlyGlyLeu-----AspValA 219
Db 770 TAGCGGAATCAGCGCGCTGTAATGAGTCCATAAGGGTGTCCGANAAGTCCATGAGTGTG 711
Qy 219 rgLysIleGlyGly-----GluAlaAspGlnValLeuAlaA 231
Db 710 GCCTGATAGTGTATTTTCCAGCAATATGCGAGGTGGAGGGTGATAGAGAGTGGGA 651
Qy 231 luThrLeuThrLysIleAlaSerGlyValGlyMet----- 242
Db 650 GTATAGGTGTGTTAGTGAATCAGGGGAAGGAGTTTGGGGTAGCGAGTGACGTAATGG 591
Qy 243 -----GluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTrpG 261
Db 590 CGAGGGCGCAGGGTTTGGCGGAGGGGATAGGGGTGATGTGGGTGGCAGGCGTCCCTTT 531
Qy 261 luIleAlaValAspAlaValGlyGlyAspProGlnAlaLeuGluLeuTrpArgGluPheG 281
Db 530 GCCTATGGTGGCGAGCGTTGGT-----GTTGAGTGGCTGACGAGATTG 486
Qy 281 luPheGlySerMetGlyArgArgAlaIleAlaIleAlaTrpSerArgGly-----LeuA 297
Db 485 GGGGAGGTAGTGTGGAATGGCTGTATTGGGGCAGTAAAGAGGGGATGATGTTGTGGGC 426
Qy 297 rg-AlaArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGluGlu 316
Db 425 GTGGCGAGGGGGGCGAGGGGC-----GCGGGGTGGTGGGGCAGGCT--- 381
Qy 317 SerAlaProValMetValAlaIlelleProAlaArgSerTrpMetMetIleArgThrCys 336
Db 380 -----CCGGGCGGTGGTGGTGGTGTACTAGAGTGTGT 348
RESULT 12
BI224374/c
LOCUS
```

1687 bp mRNA linear EST 11-JUL-2001

```
DEFINITION 602940916Fl NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104346 5',
mRNA sequence.
ACCESSION BI224374
VERSION BI224374.1 GI:14677818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1687)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E.B. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11251 row: 0 column: 03
High quality sequence stop: 195.
FEATURES
source
1..1687
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5104346"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
```

ORIGIN

Alignment Scores:

Pred. No.:	1.46	Length:	1687
Score:	119.00	Matches:	81
Percent Similarity:	33.72%	Conservative:	34
Best Local Similarity:	23.75%	Mismatches:	102
Query Match:	5.94%	Indels:	124
DB:	4	Gaps:	18

US-10-007-527A-2 (1-379) x BI224374 (1-1687)

```
Qy 61 GlyProLysGlySerGlyPheGlyGlyLeuArgSerCysGlyLysGly-----Trr 77
Db 1036 GGGAGGGTGGGAGTGGAGCGAGTGGG-----GGCTGGCGAGAGGCTGTGCGAGGTG 983
Qy 78 lIleCys-----ProCysCysAlaGlyLys-----ValGlyAlaHisArg 90
Db 982 CTGTGTATGACGCGGGCAGTGTGTGCGCGAGGCGCGCGGAGTGGGTGCGCGCG 923
Qy 91 AlaAspGluIleSerGlnValValAlaHisGlnLeuGlyThrGlySerValAlaMetVal 110
Db 922 GCGCGAGAGGTG---AAGGGAGGGCGAGCCGAGGGGTAGAGGAGCG----- 878
Qy 111 ThrMetThrMetArgHisThrAlaGlyGlnArg---LeuHisAspLeuThrThrGlyLeu 129
Db 877 -----GCTGGAGCACGCTGAGGAGAGCTGTGTGTGGGGGGGGCTG 839
Qy 130 SerAlaAla-TrrLysAlaAlaThrAsnGlyArgArgTrpArgThrGluArgGluMetTy 149
Db 838 ACATGAGCAGTGGCGGAGGGCGGAGAGCGGTGCCCTGGAGGAGC----- 793
Qy 149 rgCysAspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGlyTrrHisVa 169
Db 792 -----AGGGCGGTGGCGA--- 781
```

```

QY 169 lHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuCluSerPh 189
Db 780 -----CATGCT----- 775
QY 189 eSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAlaPr 209
Db 774 -----GGCAACACGCGGC 761
QY 209 oLeuArgAnSerGlyGly----- 215
Db 760 GCGCAGCAACGCGCGGAGAGCGCGCGGAGAGGAGAGAGAGAGAGAGAGGTGG 701
QY 216 -----LeuAspValArgLysIleGlyGlyGlu-----AlaAspGlnValLeuAl 230
Db 700 GGAGCCAGTGTGGGACCGCGCGCAGCGCGGCGGAGAGGTGGGTGACAAAGTGTCCGCG 641
QY 230 aLaTyrLeuThrLysIleAlaSerGlyValGlyMetGluValGlySer----- 246
Db 640 G-----GAGAGCGCGCGCGGTGTGAGCAGCGGAGAGCCCGTGGGAGAA 599
QY 247 ----GlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyrGluIleAlaValAs 265
Db 598 GATGGTGGGCGGAAATGATGAGACACGCGAGCGGAGCGTGGCGGAGCGGAGGTG 539
QY 265 pAlaVal-----GlyGlyAspProGlnAlaLeuGluLeuTyrArgGluPh 280
Db 538 GGTGACGAGAGCTGTGCGCAGAGGGGGCAACGGCGGAA-----TGGAGCGCGG 488
QY 280 eGluPheGlySerMetGlyArgArgAlaIleAlaTrp-----SerArgGlyLeuArgAl 298
Db 487 GCGCGCGGGGCTGAGGAAAGCGGGGCGAGATGCTGTCAGCGCAGCAATGTTGTGCGCG 428
QY 298 aArgAlaGlyLeuGlyAlaGluLeuThrAspAlaGlnIleValGluGlnGluSerA 318
Db 427 AGCTAGGCGCATGTCGCGGAGATGCGCAGCGCGGAGCAGCGAGCGGCGGAGAG 368
QY 318 laProValMetValAlaIleIle-----ProAlaArgSerTrpMetIleArgT 335
Db 367 CGCGCTGGGCGGAGGCGGAGGCGGAGGCTGAGCTGCGCAGCGGCGGAGTCTGTGCGCG 308
QY 335 hrCysAlaProTyrValPheGlyGlyLeuLeuValGluAlaGlyAlaThrTrp 354
Db 307 TGTGACGCGCAGCGGTGCCAGCAGGTGTGTGCTGCTCCATGGGMAATTAGAGCTGTGG 249

```

```

RESULT 13
BQ714786/c 851 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8482181 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6307476
DEFINITION 5', mRNA sequence.
ACCESSION BQ714786
VERSION BQ714786.1 GI:21853685
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

```

```

ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: Susan L. Sullivan, PhD.
cDNA Library Preparation: ResGen, Invitrogen Corp
DNA Sequencing By: the I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13725 row: a column: 13
High quality sequence stop: 666.
Location/Qualifiers
1. .851
FEATURES
source

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6307476"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 129"
/notes="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 2.2 kb. Constructed
by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.705 Length: 851
Score: 118.00 Matches: 78
Percent Similarity: 32.43% Conservative: 30
Best Local Similarity: 23.42% Mismatches: 112
Query Match: 5.89% Indels: 113
DB: 5 Gaps: 15

US-10-007-527A-2 (1-379) x BQ714786 (1-851)
QY 68 GlyGlyLeuArgSerCysGlyLysGlyTyrPileCysProCysCysAlaGlyLysValGly 87
Db 845 GCGGGAACCGATACGACCGTGGGGATAC-----ACCGTTTGATTTG 801
QY 88 AlaHisArgAlaAspGluIleSerGlnValAlaAlaHisGlnLeu----- 102
Db 800 GCGCCATGGCG-----CTCTCAAAGCGTTCGCGCATCATCTCTCAACCCCTTGCTT 747
QY 103 -----GlyThrGlySerValAlaMetValThrMetThrMetArgHisThr----- 117
Db 746 GACTTTGGCAAGGTGCACAGTCTCAGAGCGATGTTGGAGCATGCTCAGACATGGTGTCCAT 687
QY 118 AlaGlyGlnArgLeu-----HisAspLeuTyrThrGlyLeuSerAlaAlaTrpLysAla 135
Db 686 GTGCGGACATCTCAACTAGCACCTGATCTCTAGGTAGGTAGCAGACGCTGGGAGACC 627
QY 136 -----AlaThrAsnGlyArgArgTrpArgThrGluArgGluMet 148
Db 626 AGGCTGAGCGCAGCATGGAAGCGGTGGTAGCAGGAGCTTCTCGACACATGGTTGCTC 567
QY 149 TyrGlyCysAspGlyTyrValArgAlaValGluIleThrHisGlyLysAsnGlyTrpHis 168
Db 566 TGTGGCACCTGGGCATC-----TGGGT 543
QY 169 ValHisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSer 188
Db 542 ACTAGGCTCCACAGCATCTCTTGGACTCGGC----- 510
QY 189 PheSerAspAlaMetPheAspArgTrpThrSerLysLeuValSerLeuGlyPheAlaAla 208
Db 509 -----TCTCAGGCGCGTGGCTGGGCTCCCAACTT 480
QY 209 ProLeuArgAsnSerGlyGlyLeuAspValArgLysIleGlyGlyGluAlaAspGlnVal 228
Db 479 CCGTCCCGCTCAGCGAGGACTGGAGCTCAGA-----GACGAGATC 438
QY 229 LeuAlaAlaTyrLeuThrLysIleAlaSerGlyVal---GlyMetGluValGlySerGly 247
Db 437 CTCTGAGGGATCTTCCCGGCTGTGACCAACATAACAGTTGACTCTTTGTTGCTT 378
QY 248 AspGlyLysSerGlyArgHisGlyAsnArgAlaProTyrGluIleAlaValAspAlaVal 267
Db 377 GACGCGGATGGGCAACGCGGCGGCTCAGCTGAGGGCTAGGCTCAGC----- 324
QY 268 GlyGlyAspProGlnAlaLeuLeuTyrArgGluPheGluPheGlySerMetGlyArg 287
Db 323 -----TGG-----AGG 318
QY 288 ArgAlaIleAlaTrpSerArgGlyLeuArgAlaGlyLeuGlyAlaGluLeuThr 307

```

```

Db 317 GCTGGGCTCAGCTGGGAGACTGGCTCAGGCAGAGAACTGGGGCAGGA----- 270
Qy 308 AspaAlaGlnIleValGluGlnGluSerAlaProValMetValAlaIleProAla 327
Db 269 -----GTCAATCCAGGAGCGGTAACCT-----CCGGCG 240
Qy 328 ArgSerTrpMetMetIleArgThrCysAlaProTyrValPheGlyGluIleLeuGlyLeu 347
Db 239 CGC-----GAGCGCGCGGGCTG 222
Qy 348 ValGluAlaGlyAlaThrTrp-GluAsnLeuArgAspHisLeuHisTyrArgLeuProAl 367
Db 221 GTCACTCCGGGTTTCGGCTGGGGCCGGCGGCTTGGCGGTCTGCTCCGGGTGAAGCG 162
Qy 367 aAlaAspValArgProProlIleIleSerValArglys 379
Db 161 AGGAGCCACGAGGCCACCCAGGCTCGGGGTGAGACAG 125

RESULT 14
BZ554247/c
LOCUS
DEFINITION
pacsl-60_4590.y2 pacsl-60 Pseudomonas aeruginosa genomic clone
pacsl-60_4590, genomic survey sequence.
ACCESSION
BZ554247.1 GI:27161340
VERSION
BZ554247.1
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 893)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
Location/Qualifiers
1..893
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacsl-60_4590"
/clone_lib="pacsl-60"
/note="clinical isolate 1-60 Whole genomic shotgun
library."
ORIGIN
Alignment Scores:
Pred. No.: 0.754 Length: 893
Score: 118.00 Matches: 74
Percent Similarity: 34.07% Conservative: 19
Best Local Similarity: 27.11% Mismatches: 116
Query Match: 5.89% Indels: 66
DB: 8 Gaps: 10

US-10-007-527A-2 (1-379) x BZ554247 (1-893)

Qy 16 ProValLeuValSerSerAspLysArgGlyLeuArgHisGluLeuArgProLysLeuGln 35
Db 730 CCAGTCTGTCTCCGGCGAAACCCCTTCGGGGGTTTCAATTCCGGGAC-AGCCAGAGCCGACG 672
Qy 36 GlnIleThrThrSer-----GluThrPheAsnAlaCysGlyArgProIle 50
Db 671 CAGGCGACCCCGACATGTCGCGCCAGTGTCTCGGCTTTTCAGCGCGGTGAACGGCGCGG 612

```

```

Qy 51 SerGlyValAsnGlyValThrIleValAsnGlyProLysGlySerGlyPheGlyGlyLeu 70
Db 611 GTTGGC-----GTAGNCCCGAGCAGCGCGCGCGCGCGAGCGGGTGC 570
Qy 71 ArgSerCysGlyLysGlyTyrIleCysProCysCysAlaGlyLysValGlyAlaHisArg 90
Db 569 CAGNCT-----GCCGTCGCAATTGGAGACCACTGC 540
Qy 91 AlaAspGluIleSerGlnValAlaHisGlnLeuGlyThrGlySerValAlaMetVal 110
Db 539 GTTCAGGGTATCCAGGTCGTAGAGCATCCAGCCGAGCAGGC---GTTGTAGAAGTC 483
Qy 111 ThrMetThrMetArgHisThrAlaGlyGlnArgLeuHisAspLeuTyrThrGlyLeuSer 130
Db 482 GATCAGACCCGCCAGGTGCGGCGCTGCAACAGCA-CGTTTCGCGGACAGGTTCGGCAT 424
Qy 131 AlaAlaTyrLysAlaAlaThrAsnGlyArgArgTyrArgThrGluArgGluMetTyrGly 150
Db 423 GCAGTTGGCGCGCGGAGCCGCGCGCTCGGCATCCAGCGCGCGCATCTCCGCCAGGG 364
Qy 151 CysAspGlyTyrValArgAla---ValGluIleThrHisGlyLysAsnGlyTyrHisVal 169
Db 363 CCGGGCGAGCAGCGCGCGGCTGTTCCGGTAGTCGCGGGCGCA---GGTTGGCAC--- 310
Qy 170 HisValHisAlaLeuLeuMetPheSerGlyAspValSerGluAsnIleLeuGluSerPhe 189
Db 309 -----CTTGTCCAGCA 298
Qy 190 SerAspAlaMetPheAspArgTyrThrSerLysLeuValSerLeuGlyPheAla----- 207
Db 297 TCCAGGGCAGACCCAGCTGCTGGGCGCTTCGAGGATCGCGCGGGTGGCGGTGAA 238
Qy 208 -----AlaProLeuArgAsnSerGlyLeuAspValArgLysIleGly 222
Db 237 GGTGGCCGAGCAGATCGCCACTTCCTGGCAGTGTGGCGCTTCGGCTGCGCGCTCGTGGC 178
Qy 223 GlyGluAlaAspGlnValLeuAlaAlaTyrLeuThrLysIleAlaSerGlyValGlyMet 242
Db 177 GGCCGCGCAGCGCGGTTGCAGCAGCG-----CCGGCTTGCCTT 139
Qy 243 GluValGlySerGlyAspGlyLysSerGlyArgHisGlyAsnArgAlaProTyrGluIle 262
Db 138 CGAGACGGCGAGCGCTCGCGTGGCGGTGCGCAGGACGTAGGGCACCG----- 88
Qy 263 AlaValAspAlaValGlyGlyAspProGlnAlaLeuGlu 275
Db 87 -----GGGATCCACTAGTTCTAGAG 67

RESULT 15
BX898881/c
LOCUS
DEFINITION
Y28f11p5, mRNA sequence.
ACCESSION
BX898881
VERSION
BX898881.1 GI:40488387
KEYWORDS
EST.
ORIGIN
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzeae; Oryza.
1 (bases 1 to 1011)
Han, B., Feng, Q., Huang, Y.C., Ying, K., Li, Y., Guan, J.P., Zhu, J.J.,
Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L.,
Weng, Q.J., Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T.,
Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X.,
Zhang, L., Lan, L.F., Chen, W., Wu, S.A. and Xue, Y.B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences

```

Search completed: October 23, 2004, 06:48:11
Job time : 3882 secs